

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:11:58 ; Search time 12.48 seconds
(without alignments)
200.149 Million cell updates/sec

Title: US-09-662-784-6
Sequence: 1 DFMRCALVLAIVTQALGV.....CMGEAVONTVEKLTLGR 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	552	100.0	111	US-07-662-193-4	Sequence 4, App1
2	552	100.0	111	US-08-300-928C-6	Sequence 6, App1
3	552	100.0	111	US-08-430-944D-6	Sequence 6, App1
4	552	100.0	111	US-08-430-944D-6	Sequence 6, App1
5	552	100.0	111	US-08-430-944D-6	Sequence 6, App1
6	541	98.0	109	US-07-807-529A-6	Sequence 6, App1
7	541	98.0	109	US-07-807-529A-6	Sequence 6, App1
8	506	91.7	109	US-07-662-193-5	Sequence 5, App1
9	506	91.7	109	US-08-300-928C-8	Sequence 8, App1
10	506	91.7	109	US-08-430-944D-8	Sequence 8, App1
11	506	91.7	109	US-08-430-944D-8	Sequence 8, App1
12	506	91.7	109	US-08-431-184-8	Sequence 8, App1
13	465	84.2	92	US-08-300-928C-13	Sequence 13, App1
14	465	84.2	92	US-08-430-944D-13	Sequence 13, App1
15	465	84.2	92	US-08-430-944D-13	Sequence 13, App1
16	465	84.2	92	US-08-431-184-13	Sequence 13, App1
17	419	75.9	90	US-08-300-928C-14	Sequence 14, App1
18	419	75.9	90	US-08-430-944D-14	Sequence 14, App1
19	419	75.9	90	US-08-430-944D-14	Sequence 14, App1
20	419	75.9	90	US-08-431-184-14	Sequence 14, App1
21	409	74.1	83	US-07-662-193-7	Sequence 7, App1
22	408	73.9	97	US-08-300-928C-10	Sequence 10, App1
23	408	73.9	97	US-08-430-944D-10	Sequence 10, App1
24	408	73.9	97	US-08-430-944D-10	Sequence 10, App1
25	408	73.9	97	US-08-431-184-10	Sequence 10, App1
26	383	69.4	83	US-07-662-193-8	Sequence 8, App1
27	370.5	67.1	82	US-08-300-928C-16	Sequence 16, App1

28	370.5	67.1	82	3	US-08-430-944D-16	Sequence 16, App1
29	370.5	67.1	82	3	US-08-430-944D-16	Sequence 16, App1
30	370.5	67.1	82	3	US-08-431-184-16	Sequence 16, App1
31	185	33.5	51	1	US-07-662-193-6	Sequence 6, App1
32	185	33.5	51	3	US-08-300-928C-15	Sequence 15, App1
33	185	33.5	51	3	US-08-430-944D-15	Sequence 15, App1
34	185	33.5	51	3	US-08-430-944D-15	Sequence 15, App1
35	185	33.5	51	3	US-08-431-184-15	Sequence 15, App1
36	128.5	23.3	96	1	US-07-807-529A-38	Sequence 39, App1
37	128.5	23.3	96	1	US-08-430-944D-103	Sequence 103, App1
38	128.5	23.3	96	3	US-08-431-184-103	Sequence 103, App1
39	128	23.2	26	1	US-07-807-529A-9	Sequence 9, App1
40	128	23.2	26	3	US-08-300-928C-19	Sequence 19, App1
41	128	23.2	26	3	US-08-430-944D-19	Sequence 19, App1
42	128	23.2	26	3	US-08-430-944D-19	Sequence 19, App1
43	128	23.2	26	3	US-08-431-184-19	Sequence 19, App1
44	128	23.2	26	5	US-08-431-184-19	Sequence 19, App1
45	118.5	21.5	27	3	US-08-300-928C-57	Sequence 57, App1

ALIGNMENTS

RESULT 1
US-07-662-193-4
Sequence 4, Application US/07662193
Patent No. 5328991
GENERAL INFORMATION:
APPLICANT: Kuo, Wei-chang
TITLE OF INVENTION: Improved Preparation of Cat Dander
TITLE OF INVENTION: Allergens for Immunotherapeutic Purposes and Uses Thereof
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,193
FILING DATE: 19910228
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: P36,207
REFERENCE/DOCKET NUMBER: IML89-02AA/IPC-0020C/IMI-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-07-662-193-4

Query Match 100.0%; Score 552; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.7e-64;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Wed May 8 09:01:33 2002

us-09-662-784-6.ra1

Page 2

OY 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
DB 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
OY 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111
DB 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111

RESULT 2
US-08-300-928C-6
Sequence 6, Application US/08300928C
Patent No. 6019972
GENERAL INFORMATION:
APPLICANT: GERTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESSES:
ADDRESSER: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.60S(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-300-928C-6

Query Match 100.0%; Score 552; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.7e-64;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
DB 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
OY 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111
DB 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111

RESULT 3
US-08-430-944D-6
Sequence 6, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.

TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
TITLE OF INVENTION: 103
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-4214
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-944D-6

Query Match 100.0%; Score 552; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.7e-64;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
DB 1 DTRGALLVLTALVTOALGVKMAETCPFIYDFEFAVANGNELLDLSLRKNATEPERTA 60
OY 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111
DB 61 MKRIQDCYVENGLISRVLDTGLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111

RESULT 4
US-08-430-014-6
Sequence 6, Application US/08430014
Patent No. 6048962
GENERAL INFORMATION:
APPLICANT: GERTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSER: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002,605 (IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-014-6

Query Match 100.0%; Score 552; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 5,7e-64;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTMGALLVLLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLTVNATEPERTA 60
DB 1 DTMGALLVLLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLTVNATEPERTA 60
QY 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCGAEAVQNTVEDLKTTLGR 111
DB 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCGAEAVQNTVEDLKTTLGR 111

RESULT 5
US-08-431-184-6
Sequence 6, Application US/08431184
Patent No. 6120769
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV3
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-431-184-6

Query Match 100.0%; Score 552; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 5,7e-64;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTMGALLVLLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLTVNATEPERTA 60
DB 1 DTMGALLVLLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLTVNATEPERTA 60
QY 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCGAEAVQNTVEDLKTTLGR 111
DB 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCGAEAVQNTVEDLKTTLGR 111

RESULT 6
US-07-807-529A-6
Sequence 6, Application US/07807529A
Patent No. 554769
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce L.
APPLICANT: Morgenstern, Jay
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Greenstein, Julia L.
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: RECOMBITOPE PEPTIDES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: One Kendall Square, Building 600
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/807,529A
FILING DATE: 19911213
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
APPLICATION NUMBER: US 07/431,565
FILING DATE: 03-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Channing, Stacey L.
REGISTRATION NUMBER: 31,095
REFERENCE/DOCKET NUMBER: IPC-027/Im1-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 494-0060
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-807-529A-6

Query Match 98.0%; Score 541; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e-62;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTAK 62
DB 1 MRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTAK 60
QY 63 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 111
DB 61 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 109

RESULT 7

PCT-US93-02462-6
Sequence 6, Application PC/FUS9302462
GENERAL INFORMATION:
APPLICANT: Gelfer, Malcolm L.
APPLICANT: Garman, Richard D.
APPLICANT: Greenstein, Julia L.
APPLICANT: Kuo, Mei-chang
APPLICANT: Briner, Thomas J.
APPLICANT: Morville, Malcolm
TITLE OF INVENTION: PEPTIDES USEFUL FOR TOLERIZATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/02462
FILING DATE: 19930325
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,116
FILING DATE: 15-JAN-1993
APPLICATION NUMBER: US 07/884,718
FILING DATE: 15-MAY-1992
APPLICATION NUMBER: 07/857,311
FILING DATE: 25-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-031PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-02462-6

Query Match 98.0%; Score 541; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.5e-62;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTAK 62
DB 1 MRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTAK 60
QY 63 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 111

DB 61 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 109

RESULT 8

US-07-662-193-5
Sequence 5, Application US/07662193
Patent No. 532891
GENERAL INFORMATION:
APPLICANT: Kuo, Mei-chang
APPLICANT: Bond, Julian
TITLE OF INVENTION: Improved Preparation of Cat Dander
TITLE OF INVENTION: Allergens for Immunotherapeutic Purposes and Uses Thereof
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,193
FILING DATE: 19910228
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: P36,207
REFERENCE/DOCKET NUMBER: IM189-02AA/IPC-002CC/IMI-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-07-662-193-5

Query Match 91.7%; Score 506; DB 1; Length 109;
Best Local Similarity 93.7%; Pred. No. 4.9e-58;
Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 DMRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTA 60
DB 1 DMRGALLVLAIVTQALGVKMAETCPFIYDFVFAVANGNELLDLSITKVNATEPERTA 60
QY 61 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 111
DB 61 KIDCCYVENGSLISRYLDGLVMTTSSSKDCMGSAVONTVEDLKINTLGR 109

RESULT 9

US-08-300-928C-8
Sequence 8, Application US/08300928C
Patent No. 6019972
GENERAL INFORMATION:
APPLICANT: Gelfer, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRPP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:

ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US(IMT-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-300-928C-8

Query Match 91.7%; Score 506; DB 3; Length 109;
Best Local Similarity 93.7%; Pred. No. 4.9e-58;
Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Oy 1 DTMRALVLTALVLTQALGVKMAETCPFYDFVFVANGNELLDLSLTKVNATEPERTA 60
Db 1 DTMRALVLTALVLTQALGVKMAETCPFYDFVFVANGNELLDLSLTKVNATEPERTA 60
Oy 61 MKKIDCVYENGLISRVLDTGLVMTTSSKDCMGEAVONTVEDLKNITLGR 111
Db 61 MKKIDCVYENGLISRVLDTGLVMTTSSKDCMGEAVONTVEDLKNITLGR 109

RESULT 10
US-08-430-944D-8
Sequence 8, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMT-044DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-944D-8

Query Match 91.7%; Score 506; DB 3; Length 109;
Best Local Similarity 93.7%; Pred. No. 4.9e-58;
Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

Oy 1 DTMRALVLTALVLTQALGVKMAETCPFYDFVFVANGNELLDLSLTKVNATEPERTA 60
Db 1 DTMRALVLTALVLTQALGVKMAETCPFYDFVFVANGNELLDLSLTKVNATEPERTA 60
Oy 61 MKKIDCVYENGLISRVLDTGLVMTTSSKDCMGEAVONTVEDLKNITLGR 111
Db 61 MKKIDCVYENGLISRVLDTGLVMTTSSKDCMGEAVONTVEDLKNITLGR 109

RESULT 11
US-08-430-014-8
Sequence 8, Application US/08430014
Patent No. 6048962
GENERAL INFORMATION:
APPLICANT: GEFTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US(IMT-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-014-8

Query Match 91.7%; Score 506; DB 3; Length 109;
Best Local Similarity 93.7%; Pred. No. 4.9e-58;
Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 1 DPMRGALLVLTALVTQALGVKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTA 60
DB 1 DPMRGALLVLTALVTQALGVKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTA 60
OY 61 MKKIDPCYVENGSLISRVLDGLVMTTSSKDCGAEAVONTVEDIKLNTLGR 111
DB 61 MKKIDPCYVENGSLISRVLDGLVMTTSSKDCGAEAVONTVEDIKLNTLGR 109

RESULT 12
US-08-431-184-8
Sequence 8, Application US/08431184
Patent No. 6120769

GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 27-Apr-1995
PRIORITY APPLICATION NUMBER: US 08/430,014
FILING DATE: 02-SEPT-1994
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US (IMI-044DV3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO.: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-431-184-8

Query Match 91.7%; Score 506; DB 3; Length 109;
Best Local Similarity 93.7%; Pred. No. 4.9e-58;
Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 1 DPMRGALLVLTALVTQALGVKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTA 60
DB 1 DPMRGALLVLTALVTQALGVKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTA 60

OY 61 MKKIDPCYVENGSLISRVLDGLVMTTSSKDCGAEAVONTVEDIKLNTLGR 111
DB 61 MKKIDPCYVENGSLISRVLDGLVMTTSSKDCGAEAVONTVEDIKLNTLGR 109

RESULT 13
US-08-300-928C-13
Sequence 13, Application US/08300928C
Patent No. 6015972

GENERAL INFORMATION:
APPLICANT: GEETTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US (IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ. ID NO.: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-300-928C-13

Query Match 84.2%; Score 465; DB 3; Length 92;
Best Local Similarity 100.0%; Pred. No. 7.7e-53;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTAMRKIDCYVENGSLISRVLD 79
DB 1 VKMAETCPFYDFVFAVANGNELLDLSITKVNATEPERTAMRKIDCYVENGSLISRVLD 60
OY 80 GLVMTTSSKDCGAEAVONTVEDIKLNTLGR 111
DB 61 GLVMTTSSKDCGAEAVONTVEDIKLNTLGR 92

RESULT 14
US-08-430-944D-13
Sequence 13, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-430-944D-13

Query Match
Best Local Similarity 100.0%; Score 465; DB 3; Length 92;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKAETCPIFYDYVFAVANGNELLDLSITKVATEPERTAMRKIDCCYVENGISRLVD 79
|||||
DB 1 VKAETCPIFYDYVFAVANGNELLDLSITKVATEPERTAMRKIDCCYVENGISRLVD 60
|||||

QY 80 GLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111
|||||
DB 61 GLVMTTSSSKDCMGAVONTVEDLKLNTLGR 92
|||||

RESULT 15
US-08-430-014-13
; Sequence 13, Application US/08430014
; Patent No. 6048962
; GENERAL INFORMATION:
; APPLICANT: GEFTER, Malcolm L. et al.
; TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
```

```

APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.605 (IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-430-014-13
```

```

Query Match
Best Local Similarity 100.0%; Score 465; DB 3; Length 92;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKAETCPIFYDYVFAVANGNELLDLSITKVATEPERTAMRKIDCCYVENGISRLVD 79
|||||
DB 1 VKAETCPIFYDYVFAVANGNELLDLSITKVATEPERTAMRKIDCCYVENGISRLVD 60
|||||

QY 80 GLVMTTSSSKDCMGAVONTVEDLKLNTLGR 111
|||||
DB 61 GLVMTTSSSKDCMGAVONTVEDLKLNTLGR 92
|||||
```

Search completed: May 7, 2002, 12:13:47
Job time: 109 sec

Wed May 8 09:01:33 2002

us-09-662-784-6.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:12:13 ; Search time 14.32 Seconds
(without alignments)
590.459 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 552

Sequence: 1 DFMGALLVLLVLTQALGV.....CMGEAVONTVEDLKLNTLGR 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	98.0	109	2	C56413
2	495	89.7	107	2	JC1127
3	88	13.9	20	2	B51283
4	76.5	13.9	2401	2	T28676
5	73	13.2	1104	2	S36773
6	72	13.0	521	1	T17196
7	72	13.0	800	2	D66712
8	71	12.9	853	2	A28668
9	71	12.9	1252	1	S77037
10	70.5	12.8	295	2	A72219
11	70.5	12.8	348	2	T00641
12	70	12.7	245	2	D82610
13	69	12.5	853	2	I54964
14	69	12.5	853	2	F85922
15	69	12.5	857	2	T37459
16	69	12.5	895	2	D72571
17	68.5	12.4	466	2	H71234
18	68.5	12.4	534	2	F96713
19	68.5	12.4	1167	2	T06146
20	68	12.3	520	1	S22399
21	67.5	12.2	346	2	C85088
22	67	12.1	587	2	A34084
23	67	12.1	1343	1	H64073
24	66.5	12.0	955	2	T39765
25	65.5	11.9	351	2	S00626
26	65.5	11.9	352	2	T05723
27	65.5	11.9	468	2	H70746
28	65.5	11.9	1302	2	T00038
29	65.5	11.9	5825	2	T12117

30	65	11.8	326	2	G86462	AIg1-like protein,
31	65	11.8	392	2	B81187	AcRf/ACRE family p
32	65	11.8	392	2	B81916	probable periplasm
33	65	11.8	522	2	B64067	Na+/H+-exchanging
34	65	11.8	530	2	C82144	Na+/H+ antiporter
35	65	11.8	870	2	G69006	DNA-directed RNA p
36	65	11.8	2412	1	J01537	genome polyprotein
37	65	11.8	4128	2	JC6306	protein kinase (EC
38	64.5	11.7	114	2	A55010	neutrophil-activat
39	64.5	11.7	229	2	F83663	serine O-acetyltra
40	64.5	11.7	833	2	T28385	ORF MSY224 probabl
41	64.5	11.7	936	2	H71728	2-oxoglutarate deh
42	64.5	11.7	1116	2	B70476	hypothetical prote
43	64.5	11.7	2526	2	T20531	hypothetical prote
44	64.5	11.7	2722	2	T20532	hypothetical prote
45	64	11.6	138	2	T07353	ribosomal protein

ALIGNMENTS

RESULT 1
C56413 major allergen Fel di chain 2 precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 01-Dec-2000
C:Accession: C56413; JC1145
R:Morgenstern, J.P.; Griffith, I.J.; Brauer, A.W.; Rogers, B.L.; Bond, J.F.; Chapman, Proc. Natl. Acad. Sci. U.S.A. 88, 9690-9694, 1991
A:Title: Amino acid sequence of Fel di, the major allergen of the domestic cat: prote
A:Reference number: A56413; MUID:92052157
A:Accession: C56413
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-109 <MOR>
A:Cross-references: GB:M77341; NID:9163822; PIDN:AAC41616.1; PID:9163823
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L. Gene 113, 263-268, 1992
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg
A:Reference number: JC1126; MUID:92241678
A:Accession: JC1145
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 18-109 <GRI>
A:Experimental source: salivary gland
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.0%; Score 541; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 7.1e-50;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRCALLVLLVLTQALGVKMAETCPFYDYVFFVANGNELLLSLTKVATPEPTAMK 62
DB 1 MRCALLVLLVLTQALGVKMAETCPFYDYVFFVANGNELLLSLTKVATPEPTAMK 60
QY 63 KIDCCYVENGLISRVLDGLVMTTISSSKDCMGEAVONTVEDLKLNTLGR 111
DB 61 KIDCCYVENGLISRVLDGLVMTTISSSKDCMGEAVONTVEDLKLNTLGR 109

RESULT 2
JC1127 major allergen chain 2 precursor, short form - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: JC1127
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L. Gene 113, 263-268, 1992
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg
A:Reference number: JC1126; MUID:92241678

A:Accession: JCI127
A:Molecule type: DNA
A:Residues: 1-107 <GRI>
A:Cross-references: GB:X62478; NID:9395406; PIDN:CAA44345.1; PID:9395407
A:Experimental source: skin
C:Genetics:
A:Gene: Ch2
A:Introns: 21/1; 81/3
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>
F:50/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match: 89.7%; Score 495; DB 2; Length 107;
Best Local Similarity: 93.6%; Pred. No. 4,8e-45;
Matches 102; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

OY 3 MRCALLVLTALVTQALSKMAETCPITFDVFRVANGNELLDLSLTKVATEPERTAMK 62
DB 1 MRCALLVLTALVTQALSKMAETCPITFDVFRVANGNELLDLSLTKVATEPERTAMK 60
OY 63 KIDPCYVENGILSRVLDGLVTTTSSSKDKCEBAVQNTVEDLKLINTLGR 111
DB 61 KIDPCYVENGILSRVLDGLV--TAINEYCMGEBAVQNTVEDLKLINTLGR 107

RESULT 3
-B53283
major cat allergen Fel d I beta chain - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 12-May-1994
C:Accession: B53283
R:Dufoort, O.A.; Carreira, J.; Niltli, G.; Polo, F.; Lombardero, M.
MOL: Immunol. 28, 301-309, 1991
A:Title: Studies on the biochemical structure of the major cat allergen Felis domesticus
A:Reference number: A53283; MUID:91287714
A:Accession: B53283
A:Status: Preliminary
A:Molecule type: Protein
A:Residues: 1-20 <DOF>

Query Match: 15.9%; Score 88; DB 2; Length 20;
Best Local Similarity: 90.0%; Pred. No. 0.0044;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 20 VKMAETCPITFDVFRVANG 39
DB 1 VKMAETCPITFDVFRVANG 20

RESULT 4
T28676
rhoprory protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28676; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
MOL: Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: Z20507; MUID:97077455
A:Accession: T28676
A:Status: Preliminary; translated from GB/EMBL/DBUJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>
A:Cross-references: EMBL:U03927; NID:g1041784; PID:g1041785; PIDN:AA841263.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
MOL: Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhoprory protein. Multiple cd
A:Reference number: A45521; MUID:91101660
A:Accession: A45521
A:Status: Preliminary

A:Molecule type: DNA
A:Residues: 2260-2401 <KEE>
A:Cross-references: GB:M34281

Query Match: 13.9%; Score 76.5; DB 2; Length 2401;
Best Local Similarity: 35.1%; Pred. No. 19;
Matches 27; Conservative 12; Mismatches 25; Indels 13; Gaps 4;

OY 34 FAVANGNELLDLSLTKV-NATEPERTAMKIDPCVE-NGILSRVLDGLVTTTSSSKD 91
DB 1732 FEENNNILDSKKKIKELNMFNAE---IKRIEDKIIENGLINLT-----IERRD 1780

OY 92 CMGEAVQNTVEDLKLNF 108
DB 1781 CMFTYKTLVETLKIKI 1797

RESULT 5
S36773
GTPase-activating protein BUD2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: cyclin ClA2; protein YKL092c; protein YKL424
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999
C:Accession: S36773; S37917; S37918; S41684; S39117; S42005; S34706
R:Paik, H.O.; Chant, J.; Herskowitz, I.
Nature 365, 269-274, 1993
A:Title: BUD2 encodes a GTPase-activating protein for Bud1/Rer1 necessary for proper
A:Reference number: S36773; MUID:93382538
A:Accession: S36773
A:Molecule type: DNA
A:Residues: 1-1104 <PAR>
A:Cross-references: EMBL:L19162; NID:g402733; PIDN:AAA34461.1; PID:g402734
R:Polh, T.M.; Polh, F.M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37897
A:Accession: S37917
A:Molecule type: DNA
A:Residues: 1-681 <POH>
A:Cross-references: EMBL:Z28092; MIPS:YKL092c
A:Experimental source: strain S288C
R:James, C.M.; Gent, M.E.; Indge, K.J.; Oliver, S.G.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37918
A:Accession: S37918
A:Molecule type: DNA
A:Residues: 237-1104 <JAM>
A:Cross-references: EMBL:Z28092; MIPS:YKL092c
A:Experimental source: strain S288C
R:Crickova, F.; Nasmyth, K.
EMBO J. 12, 5277-5286, 1993
A:Title: Yeast G(1) Cyclins CLN1 and CLN2 and a GAP-1-like protein have a role in bud f
A:Reference number: S41683; MUID:94085403
A:Accession: S41684
A:Molecule type: DNA
A:Residues: 1-436, 1-438-1104 <CVK>
A:Cross-references: EMBL:X74130; NID:9395287; PIDN:CAA52228.1; PID:9395289
R:James, C.M.
submitted to the EMBL Data Library, October 1993
A:Reference number: S39116
A:Accession: S39117
A:Molecule type: DNA
A:Residues: 237-1104 <JAZ>
A:Cross-references: EMBL:X75561; NID:g414687; PIDN:CAA53241.1; PID:g414689
R:James, C.M.; Gent, M.E.; Oliver, S.G.
Yeast 10, 257-264, 1994
A:Title: Sequence analysis of a 3.5 kb EcoRI fragment from the left arm of Saccharomy
A:Reference number: S42004; MUID:94262329
A:Accession: S42005
A:Molecule type: DNA
A:Residues: 237-1050 <JAZ>
A:Cross-references: EMBL:X75561

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:BUD2; CLA2; ERC25

A:Cross-References: SGD:S0001575; MIPS:YKL092c

A:Map position: 11L

C:Superfamily: ras-specific GAP catalytic domain homology

F:520-741/Domain: ras-specific GAP catalytic domain homology <GAP>

Query Match

Best Local Similarity 13.2%; Score 73; DB 2; Length 1104;
Best Local Similarity 23.7%; Pred. No. 18;

Matches 23; Conservative 17; Mismatches 31; Indels 26; Gaps 3;

OY 21 KMAECPYDYDFEFAVANGN---LLDLSLRKVNATEPERTAMKKIQCYYENG--- 73

DB 505 KLTQSTIIFLDIFQSLRIEEMFHVLDLAKIDGT-VSRINQKVLDSKHVNSLFRGN 563

OY 74 -----ISRVLGLVMTTSSSKDC 92

DB 564 SILTKSIEQYFFRVGNEYLSKALSALKETISNKS 600

RESULT 6

IMP dehydrogenase (EC 1.1.1.205) [similarity] - Chlorobium vibrioforme

N:Alternate names: guabl protein; IMP:NAD+ oxidoreductase; IMPDH; inosine-5'-monophospha

C:Species: Chlorobium vibrioforme

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 16-Jun-2000

C:Accession: T17196

R:Peterson, B.L.; Moeller, M.G.; Stummann, B.M.; Henningsen, K.W.

submitted to the EMBL Data Library, January 1997

A:Description: Structure and organization of a 25 kbp region of the genome of the green

A:Reference number: Z18717

A:Accession: T17196

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-521 <PEP>

A:Cross-References: EMBL:Z83933; PIDN:CAB06303.1

C:Genetics:

A:Gene: guab

A:Function: catalyzes the oxidation of inosine 5'-phosphate by NAD and one molecule C

A:Pathway: purine nucleotide biosynthesis

C:Superfamily: IMP dehydrogenase; CBS homology; IMP dehydrogenase amino-terminal homolog

C:Keywords: duplication; GMP biosynthesis; NAD; oxidoreductase; purine nucleotide biosyn

F:37-103/Domain: IMP dehydrogenase amino-terminal homology <IDHN>

F:124-177/Domain: CBS homology <CBS1>

F:238-483/Domain: CBS homology <CBS2>

F:336/Active site: Cys #status predicted

Query Match

Best Local Similarity 13.0%; Score 72; DB 1; Length 521;
Best Local Similarity 32.3%; Pred. No. 9.6;

Matches 21; Conservative 9; Mismatches 25; Indels 10; Gaps 2;

OY 54 TEPERTAMKKIDC-----YENGSLRY---LDGLVMTTSSSKDCGEAVQNTVED 103

DB 426 SEPESSDRFDASAEYTKYVPEGIEGRIPORPLDEVVYQLIGLISKMGYCGVKNIEE 485

OY 104 LKLTNT 108

DB 486 LKNT 490

RESULT 7

glycogen phosphorylase (EC 2.4.1.1) [imported] - Lactococcus lactis subsp. lactis (strai

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001

C:Accession: D86712

R:Boillot, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. in press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: A86625

A:Accession: D86712

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-800 <STO>

A:Cross-References: GB:AE005176; NID:g12723610; PIDN:AAK04798.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: glgP

C:Superfamily: phosphorylase

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 13.0%; Score 72; DB 2; Length 800;

Best Local Similarity 28.3%; Pred. No. 16;

Matches 17; Conservative 14; Mismatches 27; Indels 2; Gaps 1;

OY 31 DVEFAVANGNELLDLSLRKVNATEPERTAMKKIQCYYENGSLRYLDGLVMTTSSSK 90

DB 665 EIFEAGDGNNEV--FGLTKDEYETRRNGNRYNARDIEQNPVNRILNALIDGIVPNIK 722

RESULT 8

A28668 DNA mismatch repair protein muts - Salmonella typhimurium

C:Species: Salmonella typhimurium

C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 29-Sep-1999

C:Accession: A28668; S61500

R:Haaber, L.T.; Pang, P.P.; Sobell, D.L.; Mankovich, J.A.; Walker, G.C.

J. Bacteriol. 170, 197-202, 1988

A:Title: Nucleotide sequence of the Salmonella typhimurium muts gene required for m

A:Reference number: A28668; MUID:88086868

A:Accession: A28668

A:Molecule type: DNA

A:Residues: 1-853 <HAB>

A:Cross-References: GB:M48965; NID:g154187; PIDN:AAZ7167.1; PID:g154188

A>Note: The authors translated the codon CAG for residue 213 as Arg

R.Mills, D.M.; Bajaj, V.; Lee, C.A.

Mol. Microbiol. 15, 749-759, 1995

A:Title: A 40 kb chromosomal fragment encoding Salmonella typhimurium invasion genes

A:Reference number: S61499; MUID:95302986

A:Accession: S61500

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-853 <MIL>

A:Cross-References: EMBL:U16303; NID:g565091; PIDN:AAA80578.1; PID:g565092

C:Superfamily: DNA mismatch repair protein muts

C:Keywords: DNA binding

Query Match

Best Local Similarity 12.9%; Score 71; DB 2; Length 853;
Best Local Similarity 40.4%; Pred. No. 22;

Matches 21; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

OY 29 FVDFEFAVANGNELLDLSLRKVNATEPERTAMKKIQCYYENGSLRYLDG 80

DB 38 FLEFLYDDAKRASQDLISLTKRGASAGEPIPMAGIPHAENVYLAFLVNOG 89

RESULT 9

S77037

hypothetical protein slr0697 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S. A:Reference number: S74322; MUID:97061201

A:Accession: S77037
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1252 <RAN>
A:Cross-references: EMBL:D64005; GB:AB003339; NID:G1001779; PID:G1006579
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: hypothetical protein YK1215C

Query Match 12.9%; Score 71; DB 1; Length 1252;
Best Local Similarity 30.4%; Pred. No. 34;
Matches 24; Conservative 13; Mismatches 38; Indels 4; Gaps 2;
QY 21 KMAETCPFYDVFPAVANGNELDLSTFKVNATEPERTAMKKIQDCYVENGISRYLDG 80
DB 49 KILSENPELYD--DAVTHGIRIILGLSSNEPIA--PEKSMVMKGTVAFTALLLEOKGDP 104
QY 81 LVMTTSSSKDCMGEAVON 99
DB 105 LVMTTSGFRDGLAIAYON 123

RESULT 10
-A72219
Conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence-revision 11-Jun-1999 #text-change 21-Jul-2000
C:Accession: A72219
G:R: Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; M0ID: 99287316
A:Accession: A72219
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <ARN>
A:Cross-references: GB:AE001811; GB:AE000512; NID:G4982291; PIDN:AAD36783.1; PID:G498229
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1717
C:Superfamily: conserved hypothetical protein HI1714

Query Match 12.8%; Score 70.5; DB 2; Length 295;
Best Local Similarity 26.6%; Pred. No. 7.1;
Matches 25; Conservative 19; Mismatches 31; Indels 19; Gaps 4;
QY 12 LVYTGALGVKMAETCPFYDVFPAVANGNELDLSTFKVNATEPERTAMKKIQDCYVEN 71
DB 84 LVYT---VKMPESTYIIIDKFLVLAERNELETVMVINKMDLYDED--DLKKVRE----- 132
QY 72 GLISRYVDGLVMTTSSSKDCMGEAVONTVEDLK 105
DB 133 --LEETYSGLPIYKTSATKMG-----IEELK 158

RESULT 11
T00641
Hypothetical protein F316.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence-revision 01-Feb-1999 #text-change 08-Sep-2000
C:Accession: T00641
R:Pedersen, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Araujo,
R.; Vysotskaya, Y.S.; Tu, G.; Becker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z14197
A:Accession: T00641
A:Status: translated from GB/EMBL/DDB
A:Molecule type: DNA
A:Residues: 1-388 <FED>
A:Cross-references: EMBL:AC002396; NID:G2749918; PIDN:AAC00571.1; PID:G28865; GSPDB:GN00

A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F316.4
A:Map position: 1
A:Insertions: 46/2; 87/3; 115/2; 152/3; 169/3; 186/3; 223/1; 290/3
C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

Query Match 12.8%; Score 70.5; DB 2; Length 388;
Best Local Similarity 29.6%; Pred. No. 9.8;
Matches 24; Conservative 12; Mismatches 30; Indels 15; Gaps 3;
QY 1 DTFMGALVLLVLTGALGVKMAETCPFYDVFPAVANGNEL-----LDLSLRV 51
DB 132 EALNGRVSPLVLTGQAVSKRKEQCAHFYAV---TISEELVTFDQANSGLSLALOMA 188
QY 52 NATEPERTAMKKI---QDCYV 69
DB 189 QAKDPETAFKKIKIDGFOOCEV 209

RESULT 12
D82610
Ribose-5-phosphate isomerase A XE2015 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text-change 02-Sep-2000
C:Accession: D82610
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; M0ID: 20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: D82610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <SIM>
A:Cross-references: GB:AE004020; GB:AE003849; NID:G9107122; PIDN:AAF84817.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Brites, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Neto, E.; Docena, C.; El-Doroty, H.; Facinanci, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XE2015
C:Superfamily: Haemophilus influenzae ribose-5-phosphate isomerase

Query Match 12.7%; Score 70; DB 2; Length 245;
Best Local Similarity 27.0%; Pred. No. 6.5;
Matches 27; Conservative 13; Mismatches 22; Indels 38; Gaps 4;
QY 47 SLTKVNAETP-----BRTAMKKIQDCYVENGISRY----- 77
DB 5 TVTRVSTPAPANACAPSHIPIIINPPASAKRAAKAEI--YENDMIIGVGTSTVAY 62
QY 78 -LDGLVYT-----TSSSKDCMGEAVONTVEDLKNTLNG 110
DB 63 FIDALGRTPPKRIKGAIVSSSEOSTAHKOKHGLEVLELHTFG 102
RESULT 13
I34964
DNA mismatch repair protein [imported] - Escherichia coli

C:Species: Escherichia coli
C>Date: 07-Jun-1996 #sequence.revision 07-Jun-1996 #text_change 17-Nov-2000
C:Accession: J54964; A65054; T44995
R:Schlensog, V.; Boeck, A.
J.Bacterial. 173, 7414-7415, 1991
A>Title: The Escherichia coli fdv gene probably encodes Muts and is located at minute 58
A:Reference number: J54964; MUID:92041649
A:Accession: J54964
A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-853 <RSS>
A:Cross-references: GB:M64730; NID:g146905; PIDN:AAA24188.1; PID:g146906
R:Balthier, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A65054
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-853 <BLAT>
A:Cross-references: GB:A000357; GB:U00096; NID:g267155; PIDN:AAC75775.1; PID:g1789089;
A:Experimental source: Strain K-12, Substrains MG1655
R:Carter, P.E.; Thomson-Carter, F.M.
submitted to the EMBL Data Library, May 1998
A:Description: Comparison of the mutS-ripos region from Escherichia coli O157:H7, Verocyt
A:Reference number: 22892
A:Accession: T44995
A>Status: preliminary; translated from GB/EMBL/DBDUT
A:Molecule type: DNA
A:Residues: 1-18, 'K', '20'-336, 'E', '338-853 <CAR>
A:Cross-references: EMBL:AJ006210; PIDN:CAB43497.1
A:Experimental source: strain O157:H7
C:Genetics:
A:Gene: mutS
C:Superfamily: DNA mismatch repair protein mutS

Query Match 12.5% Score 69; DB 2; Length 853;
Best Local Similarity 40.4%; Pred. NO. 35;
Matches 21; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

OY 29 FYDVFPAVANGNELLDLSITFKVNATEPERAMRKIODCYENGSLHYVDG 80
||||| : | | | | |
Db 36 FYELPYDDAKRASQLLDISLTRKGASAGEPRIPMAPIPHAVENYLAKLVNG 87
||||| : | | | | |

RESULT 14
F85922
methyl-directed mismatch repair [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: F85922
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Diallanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85922
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-853 <STD>
A:Cross-references: GB:AE005174; NID:g12517186; PIDN:AAG57842.1; GSPDB:GN00145; UMGCP:Z4
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: mutS
C:Superfamily: DNA mismatch repair protein mutS

Gy 29 FDDVFAAANGNELLDLSLTKVNATPERPAMKKIDQCYENGISLVLG 80
 |||::|||::|||||::|::|::|::|::|
 Db 36 FYELFYDAAKRASQLDLSTKRGASAGEPIPMAGIPIHYHAENVLYAKLVNQG 87

RESULT 15
 t37459

Ribonucleotide reductase (EC 1.17.4.-), B12-dependent - Thermoplasma acidophilum
 C:Species: Thermoplasma acidophilum
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: t37459

R.Trauer, A.; Benner, S.A.
 Proc. Natl. Acad. Sci. U.S.A. 94, 53-58, 1997

A>Title: The B12-dependent ribonucleotide reductase from the archaeobacterium Thermopl.

A:Reference number: Z21699; MUID:97144393

A:Accession: t37459

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-657 <TAU>

A:Cross-references: EMBL:U073619; PIDN:AAB18239.1

A>Note: cofactor vitamin B12

C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase large chain

C:Keywords: oxidoreductase

Query Match	12.5%	Score 69;	DB 2;	length 857;
Best local Similarity	27.4%	Pred. No. 36;		
Matches	26;	Conservative	17;	Mismatches 26; Gaps 5
QY	30	YDVFVAVANGN----	ELLDLSTKYNATPEPRTAKKTIODCVENGILSRVLDGLVMTT	85
		: :	: : : : : : : : :	
Db	25	YRAMTSVKNGTKMADEQLADKVAARIK--	DKEPVSVEETOD--VVEDVLMTMSKIDGTFTD	81
QY	86	ISSS-----	KDCGGEAVQNTVEDIKL	106
	: :	:		
Db	82	VAKSTYLIREKKRAIRREEKELMG----	VKDDIKL	111

Search completed: May 7, 2002, 12:14:09
Job time: 116 sec

Wed May 8 09:01:34 2002

us-09-662-784-6.rpr

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May '7, 2002, 12:13:48 ; Search time 11.68 Seconds

(without alignments)
348.441 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 552

Sequence: 1 DTRGALVLLVLTQALGV.....CMGEAVONTVEDIKLNTLGR 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	98.0	109	1	FEL2_FELCA
2	78	14.1	1247	1	IF3X_CAEEL
3	73	13.2	1104	1	IMDH_YEAST
4	72	13.0	521	1	IMDH_CHLVI
5	71	12.9	854	1	MUTS_SALTY
6	69	12.5	853	1	MUTS_ECO57
7	69	12.5	853	1	MUTS_ECOLI
8	68.5	12.4	375	1	MPAL_CHAOB
9	68	12.3	520	1	AMP_LARATH
10	67	12.1	587	1	BARL_YEAST
11	67	12.1	1343	1	RPOB_HAEIN
12	65.5	11.9	351	1	PERB_ARMRU
13	65.5	11.9	468	1	HEMI_MYCTU
14	65.5	11.9	1310	1	ACNL_HUMAN
15	65	11.8	522	1	NHAB_HAEIN
16	65	11.8	870	1	RPA1_METH
17	65	11.8	1390	1	RPOB_MTCGA
18	65	11.8	2412	1	POL1_BAYMG
19	65	11.8	4128	1	PRKD_MOUSE
20	64.5	11.7	114	1	SZ05_HUMAN
21	64.5	11.7	936	1	ODO1_RICPR
22	64.5	11.7	1116	1	YK54_AQUAE
23	64	11.6	138	1	R88_CHLVU
24	64	11.6	355	1	C3X1_HUMAN
25	64	11.6	481	1	AGM1_ARATH
26	64	11.6	556	1	ACM1_ARATH
27	63.5	11.5	319	1	ERA_TREPA
28	63.5	11.5	351	1	HRBC_RAT
29	63.5	11.5	416	1	PGK_PLAFA
30	63.5	11.5	1018	1	SCA4_RICUA
31	63	11.4	132	1	R88_CYAPA
32	63	11.4	287	1	DMSG_ECOLI
33	63	11.4	287	1	NAPH_ECOLI

ALIGNMENTS

RESULT ID	1	FEL2_FELCA	STANDARD:	PRT:	109 AA.
AC	P30440:	FEL2_FELCA			
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DE	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	MAJOR ALLERGEN I POLYPEPTIDE CHAIN 2 PRECURSOR (ALLERGEN FEL D 1-B)				
DE	(FEL D 1-B) (ALLERGEN CAT-1) (AG4) (FD1).				
GN	CH2.				
OS	Felis silvestris catus (Cat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.				
OX	NCBI_TaxID=9685;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 18-100.				
RX	MEDLINE=92052157; PubMed=1946388;				
RA	Morgenstern J.P., Griffith I.J., Brauer A.W., Rogers B.L.,				
RA	Bond J.F., Chapman M.D., Kuo M.-C.;				
RT	"Amino acid sequence of Fel di, the major allergen of the domestic				
RT	cat: protein sequence analysis and cDNA cloning."				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=92241678; PubMed=1572548;				
RA	Griffith I.J., Craig S., Pollock J., Yu X.-B., Morgenstern J.P.,				
RA	Rogers B.L.;				
RT	"Expression and genomic structure of the genes encoding Fd1, the				
RT	major allergen from the domestic cat."				
RL	Gene 113:263-268(1992).				
RN	[3]				
RP	SEQUENCE OF 18-37, AND CHARACTERIZATION.				
RX	MEDLINE=91287714; PubMed=1712068;				
RA	Duffort O.A., Carreira J., Nitti G., Polo F., Lombardero M.;				
RT	"Studies on the biochemical structure of the major cat allergen Felis				
RT	domesticus I."				
RL	Mol. Immunol. 28:301-309(1991).				
RN	[4]				
RP	CHARACTERIZATION.				
RX	MEDLINE=84265679; PubMed=6747135;				
RA	Leitman K., Ohman J.L. Jr.;				
RT	"Cat allergen 1: Biochemical, antigenic, and allergenic properties."				
RL	J. Allergy Clin. Immunol. 74:147-153(1984).				
CC	- SUBUNIT: HETEROTRIMER COMPOSED OF TWO NON-COVALENTLY LINKED				
CC	DISULFIDE-LINKED HETERODIMER OF CHAINS 1 AND 2.				
CC	- ALTERNATIVE PRODUCTS: THE LONG (CH2L) AND THE SHORT (CH2S) FORMS				
CC	MAY ARISE BY ALTERNATIVE SPLICING OR MAY REPRESENT DIFFERENT				
CC	ALILES OF THE CH2 GENE. THE SEQUENCE SHOWN HERE IS THAT OF THE				
CC	LONG FORM (CH2L).				
CC	- TISSUE SPECIFICITY: THE LONG FORM IS PREFERENTIALLY EXPRESSED IN				
CC	THE SALIVARY GLAND, WHILE THE SHORT FORM IS PREFERENTIALLY				
CC	EXPRESSED IN THE SKIN.				
CC	- DISEASE: MAJOR ALLERGEN PRODUCED BY THE DOMESTIC CAT.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M77341; AAC41616.1; -
 DR EMBL: X62478; CAA44345.1; -
 DR PIR: JG1127; JG1127.
 DR PIR: JG1145; JG1145.
 KW Allergen: Glycoprotein; Signal; Polymorphism; Alternative splicing.
 FT SIGNAL 1 17
 FT CHAIN 18 109
 FT CARBOHYD 50 50
 FT VARSPPLIC 82 82
 FT VARSPPLIC 83 109
 FT VARIANT 72 72
 FT VARIANT 72 72
 FT VARIANT 74 75
 FT VARIANT 82 83
 FT VARIANT 85 85
 FT VARIANT 86 86
 FT VARIANT 86 86
 FT VARIANT 87 88
 FT VARIANT 89 89
 FT VARIANT 96 96
 FT VARIANT 105 105
 FT CONFLICT 24 24
 FT CONFLICT 32 32
 SQ SEQUENCE 109 AA: 11854 MW: 857B9CD76036CB9 CRC64:

Query Match 98.0%; Score 541; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 2.9e-50;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MRGALLVLTALVTQALGYKMAETCPFYDFVFAVANGNELLDLSTFKVNAPEPTAMK 62
 DB 1 MRGALLVLTALVTQALGYKMAETCPFYDFVFAVANGNELLDLSTFKVNAPEPTAMK 60
 OY 63 KIDCCYVENCISRVLDGIWTTTSSKDCGKGAONVDEKLKNTLGR 111
 DB 61 KIDCCYVENCISRVLDGIWTTTSSKDCGKGAONVDEKLKNTLGR 109

RESULT 2
 ID IF3X_CAEEL STANDARD: PRT: 1247 AA.
 AC P34466;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT (EIF-3).
 GN F55H2.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
 RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,

RA Woldman P.;
 RT "2.2 kb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 CC -1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
 CC METHIONYL-TRNAI AND MRNA (BY SIMILARITY).
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EIF-3 P135 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z27080; CAA81605.1; -
 DR EMBL: Z22176; CAA81605.1; JOINED.
 DR EMBL: Z22176; CAA80143.1; JOINED.
 DR EMBL: Z27080; CAA80143.1; JOINED.
 KW Wormpep: F55H2.6; GE00213.
 KW Hypothetical protein; Initiation factor; Protein biosynthesis.
 SQ SEQUENCE 1247 AA: 139920 MW: C64F3EF2708774F CRC64:

Query Match 14.1%; Score 78; DB 1; Length 1247;
 Best Local Similarity 28.0%; Pred. No. 2.9;
 Matches 26; Conservative 14; Mismatches 37; Indels 16; Gaps 2;

OY 28 IFYDFVFAVANGNELLDLSTFKVNAPEPTAMK KIDCCYVENCIL-----SRV 77
 DB 384 IADYVNAARGVOSILDGNILAINPGEDKTH-----YIMNIFSLGFDVDRHYE 437
 OY 78 LDGLVMTTSSKDCGKGAONVDEKLKNTLG 110
 DB 438 LOGDAAPAAATSTDIQVRAFAATLDDPRILNTLG 470

RESULT 3
 ID BUD2_YEAST STANDARD: PRT: 1104 AA.
 AC P33314;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INHIBITORY REGULATOR PROTEIN BUD2/CLL2.
 GN BUD2 OR CLL2 OR ERC25 OR YKL092C OR YKL424.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94085403; PubMed=8262070;
 RA Cvrckova F., Nasmyth K.;
 RT "Yeast G1 cyclins CLL1 and CLL2 and a GAP-like protein have a role in
 RT bud formation.";
 RL EMBO J. 12:5277-5286(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93382538; PubMed=8371782;
 RA Park R.-O., Chan J., Herskowitz I.;
 RT "BUD2 encodes a GTPase-activating protein for Bud1/Rsr1 necessary for
 RT proper bud-site selection in yeast.";
 RL Nature 365:269-274(1993).
 RN [3]
 RP SEQUENCE OF 1-681 FROM N.A.
 RA Pohl T.M., Pohl F.M.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 237-1104 FROM N.A.

```

RX MEDLINE=94262329; PubMed=8203166;
RA James C.M., Gent M.E., Oliver S.G.;
RT "Sequence analysis of a 3.5 kb EcoRI fragment from the left arm of
RT Saccharomyces cerevisiae chromosome XI reveals the location of the
RT MBI1 gene and a sequence related to a GTPase-activating protein.";
RL Yeast 10;257-264(1994).
CC -1- FUNCTION: STIMULATES THE GTPASE ACTIVITY OF BUD1/RSR1.
CC PARTICIPATES IN THE REGULATION OF BUD-SITE SELECTION.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RAS-GAP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announcement/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL; X74130; CAA52228.1; -;
DR EMBL; L19162; AAA34461.1; -;
DR EMBL; X75561; CAA53241.1; -;
DR EMBL; Z28092; CAA81930.1; -;
DR PIR; S36773; S36773.
DR SGD; S0001575; BUD2.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001936; RasGAP.
DR Pfam; PF00616; RasGAP; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00323; RasGAP; 1.
DR PROSITE; PS00499; C2_DOMAIN_1; FALSE_NEG.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS00509; RAS_GTPASE_ACTIV_1; 1.
DR PROSITE; PS50018; RAS_GTPASE_ACTIV_2; 1.
KW GTPase activation.
KW DOMAIN 322 425 C2 DOMAIN.
FT DOMAIN 505 721 RAS-GAP.
FT DOMAIN 496 499 POLY-SER.
FT DOMAIN 1096 1099 POLY-LYS.
FT CONFLICT 437 437 N -> Y (IN REF. 1).
SQ SEQUENCE 1104 AA; 126662 MW; 451AFC5A78384760 CRC64;

Query Match 13.2%; Score 73; DB 1; Length 1104;
Best Local Similarity 23.7%; Pred. No. 8.4;
Matches 23; Conservative 17; Mismatches 31; Indels 26; Gaps 3;

OY 21 KMAETCPFFYDFEAVANGNE---LLDLSTKVNATEPERTAMKKIDQCYENGU----- 73
      : : | | | : : : : : : : : : : : : : : : : : : : : : :
DB 505 KLTQCSIIFLDFOSLRSIEEFHVLIDKLEAKIDGT-VSRINQKNDLSKSHVFNLSLFRGN 563
      : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 74 -----ISRVLDDLVMNTTSSSDC 92
      : : | : : : : : : : : : : : : : : : : : : : : : :
DB 564 SILKSIQDYEFFRNGEYLSKALSLKLEIESNRSK 600

RESULT 4
IMDH_CHLVI STANDARD: PRT; 521 AA.
AC 050316;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (EC 1.1.1.205) (IMP
DE DEHYDROGENASE) (IMPDH) (IMPD).
GN GUBA.
OS Chlorobium vibrioforme.
OC Bacteria; Green sulfur bacteria; Chlorobium.
OX NCBI_TaxID=1098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F, THIOSULFATOPHILUM NCIB 8327;
RA Petersen B.L., Moeller M.G., Stummann B.M., Henningsen K.W.;
```

```
RT "Clustering of genes with function in the biosynthesis of  
RT bacteriochlorophyll and heme in the green sulfur bacterium Chlorobium  
RT vibrioforme.";  
RL Hereditas 125:93-96(1996).  
CC -1- CATALYTIC ACTIVITY: INOSINE 5'-PHOSPHATE + NADH.  
CC XANTHOSINE 5'-PHOSPHATE + NADH.  
CC -1- PATHWAY: FIRST REACTION UNIOE TO GMP BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SIMILARITY: TO OTHER EUKARYOTIC AND PROKARYOTIC IMPDH AND TO  
CC GMP REDUCTASE.  
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL; Z83933; CAB06303.1; -.  
DR InterPro: IPR000644; CBS.  
DR InterPro: IPR003009; FMN enzyme.  
DR InterPro: IPR001093; IMP_DH_GMP_RED.  
DR Pfam; PF00571; CBS; 2.  
DR Pfam; PF00478; IMPDH_C; 1.  
DR Pfam; PF01574; IMPDH_N; 1.  
DR SMART; SMO0116; CBS; 2.  
DR PROSITE; PS00487; IMP_DH_GMP_RED; 1.  
KW Oxidoreductase; NAD; GMP biosynthesis; Purine biosynthesis; Repeat;  
KW CBS domain.  
RW DOMAIN 118 176 CBS 1.  
FT FT 184 237 CBS 2.  
FT BINDING 336 336 IMP (POTENTIAL).  
SQ SEQUENCE 521 AA; 56627 MW; 00655FCBCB182815a CRC64;  
  
Query Match 13.0%; Score 72; DB 1; Length 521;  
Best Local Similarity 32.3%; Pred. No. 4.5;  
Matches 21; Conservative 9; Mismatches 25; Indels 10; Gaps 2;  
  
QY 54 TEPERTAMKKIODC-----YVENGLISRV---LDGLVMTTSSSKDKCMGEAVNYVED 103  
DB :|||:::|::| |::|:|:|:|:|:|:|:|:|:|:  
Db 426 SEPESGSSTRYPQDASAEYTKKYPEGIEGRIPQRPLDEVYQLIGLGKSMGCGVKYNEE 485  
  
QY 104 LKLNKT 108  
DB ||||  
Db 486 LKKNKT 490  
  
RESULT 5  
MUTS_SALTY ID MUTS_SALTY STANDARD; PRT; 854 AA.  
AC P10339;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DNA MISMATCH REPAIR PROTEIN MUTS.  
GN MUTS.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC Salmonella.  
CX NCBI_TaxID=602;  
[1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.  
RP MEDLINE=8808868; PubMed=3275609;  
RA Haber L.T., Pang P.P., Sobell D.I., Mankovich J.A., Walker G.C.;  
RT "Nucleotide sequence of the Salmonella typhimurium muts gene required  
RT for mismatch repair: homology of Muts and Hexa of Streptococcus  
RL J. Bacteriol. 170:197-202(1988).  
[2]  
RP MUTAGENESIS, AND ATP-BINDING.
```

```

RX MEDLINE-91330898; PubMed-1651234;
RA Haber L.T., Walker G.C.;
RT "Altering the conserved nucleotide binding motif in the Salmonella
RT typhimurium Muts mismatch repair protein affects both its Atpase and
RT mismatch binding activities";
RL EMBO J. 10:2707-2715(1991).
[3]
RP IDENTIFICATION OF PROBABLE FRAMESHIFT IN 393-424.
RX MEDLINE-94087727; PubMed-7903399;
RA Claverie J.-M.;
RT "Detecting frame shifts by amino acid sequence comparison.";
RL J. Mol. Biol. 234:1140-1157(1993).
-1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
CC IN DNA. IT IS POSSIBLE THAT IT CARRY OUT THE MISMATCH RECOGNITION
CC STEP. THIS PROTEIN HAS A WEAK ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U16303; AAA80578.1; -
DR EMBL: M18965; AAA27167.1; -
DR PIR: A28668; A28668.
DR StGene: SG10237; mutS.
DR InterPro: IPR000432; Muts_C.
DR InterPro: IPR002863; Muts_N.
DR Pfam: PF00488; Muts_C.1.
DR Pfam: PF01624; Muts_N.1.
DR ProDom: PD001263; Muts_C.1.
DR SMART: SM00534; MutsC.1.
DR SMART: SM00533; MutsS.1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR DNA repair; ATP-binding; DNA-binding.
FT NP_BIND 615 622
FT MUTAGEN 621 622
FT K->A DEFECTIVE IN MISMATCH REPAIR;
FT ATPASE ACTIVITY REDUCED 6 FOLD.
FT CONFLICT 393 424
FT ALRKKKDFALRDLRLIIPYLVNDA ->
FT RCYKRWAISSPCATSWMAPLTRRWSRR (IN
FT REF. 1).
FT
SQ SEQUENCE 854 AA; 94757 MW; 322207EF90BBA36B CRC64;

Query Match 12.9%; Score 71; DB 1; Length 854;
Best Local Similarity 40.4%; Pred. No. 10;
Matches 21; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

OY 29 FYDVFAVANGNELLDISTKRVNATEPRTAMKRIQDCYVNGLSRYLDG 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 38 FYELFYDAKRAASQLDLSITKRGASAGEPIPMAGIPIHVAVENTIATLVNCG 89

RESULT 6
MUTS_ECO57 STANDARD: PRT; 853 AA.
AC ONS6P8;
DT 20-AUG-2001 (rel. 40, Created)
DT 20-AUG-2001 (rel. 40, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MUTS.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7;
RA Leclerc J.E., Li B., Payne W.L., Cebula T.A.;

RT "High mutation frequencies among Escherichia coli and Salmonella
RT pathogens.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7;
RA Carter P.E., Thomson-Carter F.M.;
RT "Comparison of the mutS-rtos region from Escherichia coli O157:H7,
RT verocytotoxin-containing and non-verocytotoxin-containing E.coli.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE-21074935; PubMed-11206551;
RA Perma N.T., Plunkett G., Iit, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfal G., Hackett J., Klink S., Boulton A., Shao Y., Miller L.,
RA Grobback E.J., Davis N.W., Lim A., Dimantia E.T., Potamotis K.,
RA Apodaca J., Anthathaman I.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Bletcher F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Okunishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL Nature Res. 8:11-22(2001).
-1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
CC IN DNA. IT IS POSSIBLE THAT IT CARRY OUT THE MISMATCH RECOGNITION
CC STEP. THIS PROTEIN HAS A WEAK ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U69873; AAB97931.1; -
DR EMBL: AJ006210; CAB43497.1; -
DR EMBL: AE005501; AAG57842.1; -
DR EMBL: AP002562; BAB37012.1; -
DR InterPro: IPR000432; Muts_C.
DR InterPro: IPR002863; Muts_N.
DR Pfam: PF00488; Muts_C.1.
DR Pfam: PF01624; Muts_N.1.
DR ProDom: PD001263; Muts_C.1.
DR SMART: SM00534; MutsC.1.
DR SMART: SM00533; MutsS.1.
DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
DR DNA repair; ATP-binding; DNA-binding.
FT NP_BIND 614 621
FT CONFLICT 59 59
FT CONFLICT 144 144
FT CONFLICT 317 317
FT G -> S (IN REF. 1).
FT G -> A (IN REF. 1).
FT D -> H (IN REF. 1).
SQ SEQUENCE 853 AA; 95290 MW; 9D6BB8207F9F87AB CRC64;

Query Match 12.5%; Score 69; DB 1; Length 853;
Best Local Similarity 40.4%; Pred. No. 17;
Matches 21; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

OY 29 FYDVFAVANGNELLDISTKRVNATEPRTAMKRIQDCYVNGLSRYLDG 80
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 36 FYELFYDAKRAASQLDLSITKRGASAGEPIPMAGIPIHVAVENTIATLVNCG 87
```

RESULT 7
MUTS_ECOLI STANDARD: PRT: 853 AA.
AC P23909; P71279;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DNA MISMATCH REPAIR PROTEIN MUTS.
GN MUTS OR FDV OR B2733.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92041649; PubMed=1938937;
RA Schlenso V., Boeck A.;
RT "The Escherichia coli fdv gene probably encodes muts and is located
RT at minute 58.8 adjacent to the hyc-hyp gene cluster.";
RL J. Bacteriol. 173:7414-7415(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 602-731 FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Denamur E.;
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
CC IN DNA. IT IS POSSIBLE THAT IT CARRY OUT THE MISMATCH RECOGNITION
CC STEP. THIS PROTEIN HAS A WEAK ATPASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64730; AAA24188.1; -
DR EMBL: U29579; AAA69243.1; -
DR EMBL: AE000357; AAC7575.1; -
DR EMBL: AF004287; AAD0197.1; -
DR EMBL: AF001987; AAD00921.1; -
DR EMBL: AF001988; AAD00922.1; -
DR EMBL: AF001989; AAD00923.1; -
DR EMBL: AF001990; AAD00924.1; -
DR EMBL: AF001991; AAD00925.1; -
DR EMBL: AF001992; AAD00926.1; -
DR EMBL: AF001993; AAD00927.1; -
DR EMBL: AF001994; AAD00928.1; -
DR EMBL: AF001995; AAD00929.1; -
DR EMBL: AF001996; AAD00930.1; -
DR EMBL: AF001997; AAD00931.1; -
DR EMBL: AF001998; AAD00932.1; -
DR EMBL: AF001999; AAD00933.1; -
DR EMBL: AF002000; AAD00934.1; -
DR EMBL: AF002001; AAD00935.1; -
DR EMBL: AF002002; AAD00936.1; -
DR EMBL: AF002003; AAD00937.1; -
DR EMBL: AF002004; AAD00938.1; -
DR EMBL: AF002005; AAD00939.1; -

DR EMBL: AF002006; AAD00940.1; -
DR EMBL: AF002007; AAD00941.1; -
DR EMBL: AF002008; AAD00942.1; -
DR EMBL: AF002009; AAD00943.1; -
DR EMBL: AF002010; AAD00944.1; -
DR SWISS-2DPAGE; P23909; COLI.
DR ECODBASE; E093.0; 6TH EDITION.
DR Ecogene; EGI0625; muts.
DR InterPro; IPR000432; MUTS_C.
DR InterPro; IPR002863; MUTS_N.
DR Pfam; PF00488; MUTS_C; 1.
DR Pfam; PF01624; MUTS_N; 1.
DR ProDom; PD001263; MUTS_C; 1.
DR SMART; SM00533; MUTSd; 1.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA repair; ATP-binding; DNA-binding; Complete proteome.
FT NP_BIND 614 621 ATP (POTENTIAL).
SQ SEQUENCE 853 AA; 95246 MW; 5A1F295362671D55 CRC64;

Query Match 12.5%; Score 69; DB 1; Length 853;
Best Local Similarity 40.4%; Pred. No. 17;
Matches 21; Conservative 5; Mismatches 26; Indels 0; Gaps 0;

QY 29 FVDVFFAVANGNELLDLSLFVNATEPERFTAMKRIQDCYVNGLSRVLDG 80
DB 36 FTELFTDDAKRASOLDLSLIRGASAGEPIPMAGIPYHAYENYAKLVNOG 87

RESULT 8
MPAL_CHAOB STANDARD: PRT: 375 AA.
AC Q96385;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR POLLEN ALLERGEN CHA O 1 PRECURSOR.
OS Chamaecyparis obtusa (Japanese cypress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC Chamaecyparis.
OX NCBI_TaxID=13415;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Pollen.
RX MEDLINE=96265194; PubMed=8676896;
RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
RA Ohta N.;
RT "Purification, characterization and molecular cloning of Cha o 1, a
RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
RL Mol. Immunol. 33:451-460(1996).
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D45404; BA08246.1; -
DR Mendel; 7626; ChaoB; 1088; 7626.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 375 MAJOR POLLEN ALLERGEN CHA O 1.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 375 AA: 40258 MW: 81CD91DF066DBF CRC64;

Query Match 12.48; Score 68.5; DB 1; Length 375;
Best Local Similarity 27.08; Pred. No. 7.3;
Matches 17; Conservative 18; Mismatches 23; Indels 5; Gaps 2;

OY 38 NGNELLIDS-LTKVATEPERFAMKRIKIDCYENGLISRVLDGLVTTISSKDCMGEA 96
DB 152 SGNVLISEASGVVPHADGDAITMRVYDWMIDHNSLSDSDGLVDVYLAFT---GVT 207
OY 97 VON 99
DB 208 ISN 210

RESULT 9

AMPL_ARATH STANDARD; PRT: 520 AA.

AC P30164;
DT 01-APR-1993 (rel. 25, Created)
DT 01-APR-1993 (rel. 25, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE CYTOSOL AMINOPEPTIDASE (EC 3.4.11.1) (LEUCINE AMINOPEPTIDASE) (LAP)
DE (LEUCYL AMINOPEPTIDASE) (PROLINE AMINOPEPTIDASE) (EC 3.4.11.5) (PROLYL
DE AMINOPEPTIDASE).
GN ATG24200 OR F27D4.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RP SEQUENCE FROM N. A.
RC STRAIN=CV LANDSBERG REECTA; TISSUE=Leaf;
RX MEDLINE=92209533; PubMed=1555602;

RA Bartling D., Weller E.W.;
RT "Leucine aminopeptidase from Arabidopsis thaliana. Molecular evidence
for a phylogenetically conserved enzyme of protein turnover in higher
plants.";
RL Eur. J. Biochem. 205:425-431(1992).

RN [2]
RP SEQUENCE FROM N. A.
RC STRAIN=CV. COLUMBIA;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Vanaken S.E., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cressy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Press D.,
RA Nieman W.C., White O., Eisele J.A., Salberg S.L., Fraser C.M.,
RA Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).

CC -1- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
UNSUBSTITUTED AMINO-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES.
CC -1- CATALYTIC ACTIVITY: RELEASE OF AN N-TERMINAL AMINO ACID, XAA-1-
XBB-, IN WHICH XAA IS PREFERABLY LEU, BUT MAY BE OTHER AMINO ACIDS
CC -1- COFACTOR: BINDS TWO ZINC IONS (BY SIMILARITY).
CC -1- INCLUDING PRO ALTHOUGH NOT ARG OR LYS, BUT MAY BE OTHER AMINO ACIDS
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17; ALSO KNOWN AS THE
CYTOSOL AMINOPEPTIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.ebi.ac.uk/announcements/>
CC or send an email to license@ebi.ac.uk.)

DR EMBL: X63444; CAA5040.1; -
DR EMBL: AC005967; AAD03381.1; -
DR PIR: S22399; S22399.
DR HSSP: P00277; 1BPM.
DR MEROPS: M17.002; -
DR InterPro: IPR000819; Peptidase_M17.
DR Pfam: PF00883; Peptidase_M17; 1.
DR PRINTS: PR00481; LAMNOPEPTASE.
DR PROSITE: PS00631; CYTOSOL_AP; 1.
KW Hydrolase; Amino-peptidase; Zinc.
FT METAL 288 288 ZINC (2) (BY SIMILARITY).
FT METAL 293 293 ZINC (1 AND 2) (BY SIMILARITY).
FT METAL 313 313 ZINC (2) (BY SIMILARITY).
FT METAL 373 373 ZINC (1) (BY SIMILARITY).
FT METAL 375 375 ZINC (1 AND 2) (BY SIMILARITY).
FT ACT_SITE 300 300 POTENTIAL.
FT ACT_SITE 377 377 POTENTIAL.
SQ SEQUENCE 520 AA: 54509 MW: D3FA9CCD312AA92 CRC64;

Query Match 12.38; Score 68; DB 1; Length 520;
Best Local Similarity 28.38; Pred. No. 12;
Matches 26; Conservative 15; Mismatches 21; Indels 30; Gaps 4;

OY 46 ISLTAVNATPEPERA-----MKKIDPCYENGLISRV---L 78
DB 5 LGIDQNSTEHRHISTFAKEIDIVEMKGDILVGVTEKDLANDGSKREPDILSKVDAIL 64
OY 79 DGLVYTTISSKDCMGEAVONTVEDIKLNTLG 110
DB 65 SGL-LAOVSEDFPTKPGQSTV-LRLPLG 93

RESULT 10

BARL_YEAST STANDARD; PRT: 587 AA.

AC P12630;
DT 01-OCT-1989 (rel. 12, Created)
DT 01-OCT-1989 (rel. 12, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE BARRIERPEPSIN PRECURSOR (EC 3.4.23.35) (EXTRACELLULAR "BARRIER"
DE PROTEIN) (BAR PROTEINASE).

GN BAR1 OR SST1 OR YII015W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RP SEQUENCE FROM N. A.
RC MEDLINE=88124818; PubMed=3124102;

RX Mackay V.L., Welch S.K., Insley M.Y., Manney T.R., Holly J.,

RA Saari G.C., Parker M.L.;

RT "The Saccharomyces cerevisiae BAR1 gene encodes an exported protein
with homology to pepsin.";

RL Proc. Natl. Acad. Sci. U.S.A. 85:55-59(1988).

RN [2]
RP SEQUENCE FROM N. A.

RC STRAIN=S288C / AB972;

RA Churchill C.G., Badcock K., Bankier A.T., Bowman S., Brown D.,

RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Devlin K., Fraser A.,

RA Louis E., Lye G., Moulé S., Moulé T., Odell C., Pearson D.,

RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,

RA Walsh S.V., Whitehead S.;

RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases

CC -1- FUNCTION: THIS PROTEIN CALLED "BARRIER ACTIVITY" IS EXCRETED BY
CC YEAST CELLS MATING TYPE A. IT IS PROBABLY A PROTEASE THAT CLEAVES

CC ALPHA-FACTOR AND THUS ACTS AS AN ANTAGONIST OF THIS MATING
CC PHEROMONE AND ESTABLISHES OPTIMAL PHEROMONE CONCENTRATION FOR
CC CONUGATION.
CC -1- CATALYTIC ACTIVITY: SELECTED CLEAVAGE OF 6-LEU-1-LYS-7 BOND IN
CC THE PHEROMONE ALPHA-MATING FACTOR.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- INDUCTION: BY ALPHA-FACTOR.
CC -1- MISCELLANEOUS: IT IS FOUND ONLY IN A MATING TYPE CELLS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY
CC -1- SIMILARITY: TO OTHER EUKARYOTIC ASPARTYL PROTEASES IN THE FIRST
CC TWO THIRDS OF THE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Z46881; CAA86977.1; -
DR EMBL: J03573; AAA34451.1; -
DR PIR: A34084; A34084.
DR HSSP: P28871; IEAG.
DR MEROPS: A01.015; -
DR SGD: S0001277; BAR1.
DR InterPro: IPR001969; Asp.protease.
DR InterPro: IPR001461; Pepsin.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Pheromone response; Hydrolase; Aspartyl protease; Glycoprotein;
KV Signal.
FT SIGNAL 1 24
FT CHAIN 25 587
FT DOMAIN 25 229
FT ACT_SITE 238 396
FT ACT_SITE 63 63
FT ACT_SITE 287 287
FT DISULFID 322 358
FT CARBOHYD 84 84
FT CARBOHYD 90 90
FT CARBOHYD 268 268
FT CARBOHYD 308 308
FT CARBOHYD 366 366
FT CARBOHYD 398 398
FT CARBOHYD 468 468
FT CARBOHYD 503 503
FT CARBOHYD 551 551
SQ SEQUENCE 587 AA; 63729 MW; CC21DB7FDBC83984 CRC64;
Query Match 12.1%; Score 67; DB 1; Length 587;
Best Local Similarity 25.4%; Pred. No. 17;
Matches 32; Conservative 20; Mismatches 50; Indels 24; Gaps 5;

AC P43738;
DT 01-NOV-1995 (rel. 32; Created)
DT 01-NOV-1995 (rel. 32; Last sequence update)
DT 20-AUG-2001 (rel. 40; Last annotation update)
DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
GN RPOB OR H10515.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95550630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utermarck T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RT Science 269:496-512(1995).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
CC RNA(N).
CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U32733; AAC22173.1; -
DR TIGR: H10515; -
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_POL_BETA; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW Transferrase; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 1343 AA; 149783 MW; 4EF99CD648686A44 CRC64;

Query Match 12.1%; Score 67; DB 1; Length 1343;
Best Local Similarity 27.3%; Pred. No. 45;
Matches 35; Conservative 20; Mismatches 37; Indels 36; Gaps 8;
OY 11 ALVYQALGVKAEPCIFED-VFPAVANGNELLDLSITKVNATEPER--TAAKKIO 65
DB 206 ATIIIRALGVYTHEILNLFEDKRTFEIA-GDKRLMLT-----VPERLRGRTAFSDIE 256
OY 66 ---DCYVENG--LISRYLDGLVYTTIS-----SSKDCW---GEAVQNTVED 103
DB 257 ANGKVVYVERGRITARRHAKLEKDNISQVVPSEYIIIGRVASKDYDLESGETICPANGR 316
OY 104 LKNTLTGR 111
DB 317 ISLETLAK 324
RESULT 12

PERB_ARXRU STANDARD: PRT: 351 AA.
ID PERB_ARXRU
AC P15232;
BT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE PEROXIDASE C1B PRECURSOR (EC 1.11.1.7).
GN PRX1B OR HPRC2.
OS Amoracia rusticana (Horse radish) (Amoracia laphatfolia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Amoracia.
OX NCBI_TaxID=3704;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=68225087; PubMed=3371352;
RA Fujiyama K., Takemura H., Shibayama S., Kobayashi K., Choi J.K.,
RA Shimizu A., Takano M., Yamada Y., Okada H.,
RL Eur. J. Biochem. 173:681-687(1988). OXIDATION OF TOXIC REDUCTANTS,
CC HIOSYNTHEIS AND DEGRADATION OF LIGNIN. DEFENSE RESPONSE TOWARD
CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) = OXIDIZED DONOR + 2 H(2)O.
CC -1- COFACTOR: HEME.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M37157; AAA33378.1; -;
DR PIR: S00626; S00626.
DR HSSP: P00433; 2ATJ3.
DR InterPro: IPR002016; Peroxidase.
DR Pfam: PF00141; Peroxidase_1.
DR PRINTS: PR00458; PEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE_1; 1.
DR PROSITE: PS00436; PEROXIDASE_2; 1.
KW Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
KW Signal.
FT SIGNAL 1 28
FT CHAIN 29 351
FT MOD_RES 29 29
FT
FT ACT_SITE 66 66 PEROXIDASE C1B.
FT ACT_SITE 70 70 PYRROLIDONE CARBOXYLIC ACID (BY
FT ACT_SITE 198 198 SIMILARITY).
FT DISULFID 39 119 DISTAL HISTIDINE (BY SIMILARITY).
FT DISULFID 72 77 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
FT DISULFID 125 329 BY SIMILARITY.
FT DISULFID 205 237 BY SIMILARITY.
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 351 AA; 38645 MW; 7A8C606A3928950B CRC64;

Query Match 11.9%; Score 65.5; DB 1; Length 351;
Best Local Similarity 26.9%; Pred. No. 14;
Matches 25; Conservative 18; Mismatches 41; Indels 9; Gaps 5;

13 LTTGACIAFYA-SLSDAQLTPEFYDT--SCPMNSVIRDTITNEL-SDPRITASILRLH 68
64 IODCVENGILSRVLDGLVMTTISSSKDCMGEA 96
69 FHDCFVNGCDASILDLDN--TSTFLTKXDALGNA 99
Db
RESULT 13
ID HEM1_MYCTU STANDARD: PRT: 468 AA.
AC 01139;
BT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE GLUTAMYL-TRNA REDUCTASE (EC 1.2.1.2) (GLUTR).
GN HEMA OR RV0509 OR MT0530 OR MTCY2063.36.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogan A., McLean S., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon M., Squares S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence".
RT Nature 393:537-544(1998).
RN
RP
RP SEQUENCE FROM N.A.
RC STRAIN-CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.N., Deboy R., Dodson R., Gwinn M.L., Hatt D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Kohnen A., Uitterlind T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains".
RT Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA(GLU) + NADPH = GLUTAMATE-1-
CC SEMIALDEHYDE + NADP(+) + TRNA(GLU).
CC -1- PATHWAY: FIRST STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC
CC EMBL: Z77162; CAB00935.1; -;
CC EMBL: AE006953; AAK44753.1; ALT_INTT.
CC TIGR: MT0530; -;
DR TubercuList: RV0509; -;
DR InterPro: IPR000343; GLUTR.
DR InterPro: IPR000594; Thif_family.
DR Pfam: PF00745; GLUTR; 1.
DR PROSITE: PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 468 AA; 49361 MW; DDC257648FF6EAAAB CRC64;

Query Match 11.9%; Score 65.5; DB 1; Length 468;

Best Local Similarity 23.3%; Pred. No. 19;
Matches 30; Conservative 18; Mismatches 52; Indels 29; Gaps 5;

OY 1 DFMGALVLLVLTQALGVKAECPICY-----DVFVAVANGNELLDLSLTKN 52
| | | | | : : : : : | | | | | : : : : :
DB 60 DAFHGLSVIGGVLAHSHSMGELTKYAYVRYSEAVHLEFVAVSG-----LD-----S 109
| | | | | : : : : : | | | | | : : : : :
OY 53 ATEPERTAMKKIQQDCVY---ENGLISRVLDGLVMTTSSSK-----DCMGEAVQNTV 101
| | | | | : : : : : | | | | | : : : : :
DB 110 AVIGEGQVVGVRRAVAVAESNRTVGRLVHLLAQLALSGKRVHSEFLDAAGASVYVA 169
| | | | | : : : : : | | | | | : : : : :
OY 102 EDIKLNTLG 110
| | | | | : : : : : | | | | | : : : : :
DB 170 LGMARKLIG 178
| | | | | : : : : : | | | | | : : : : :

RESULT 14
ACNL_HUMAN STANDARD; PRT; 1310 AA.
ID ACNL_HUMAN 014525; 060799;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ASTROTACTIN 1 (FRAGMENT).
GN ASTN1 OR ASTN OR KIA0289.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA MEDLINE=97323006; PubMed=9179496;
RA Ohara O., Nagase T., Ishikawa K.-I., Nakajima D., Ohira M.,
RA Seki N., Komura N.;
RT Construction and characterization of human brain cDNA libraries
RT suitable for analysis of cDNA clones encoding relatively large
RT proteins.";
RT DNA Res. 4:53-59(1997).
RN [2]
RP SEQUENCE OF 166-906 FROM N.A.
RA Pearce A.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEURONAL ADHESION MOLECULE THAT IS REQUIRED FOR GLIAL-
CC GUIDED MIGRATION OF YOUNG POSTMITOTIC NEUROBLASTS IN CORTICAL
CC REGIONS OF DEVELOPING BRAIN, INCLUDING CEREBRUM, HIPPOCAMPUS,
CC CEREBELLUM AND OLFACTORY BULB.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 AND 2 (SHOWN HERE); MAY BE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB006627; BAA22958.1; -;
DR EMBL; AL022145; CA18140.1; -;
DR EMBL; AL022145; CA18141.1; -;
DR MIM: 600904; -;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001777; FN-III.
DR InterPro: IPR01862; MAC_perforin.
DR Pfam: PF00041; fn3; 1.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00001; EGF-like; 1.
DR SMART: SM00060; FN3; 1.
DR SMART: SM00457; MACPF; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.

DR PROSITE: PS01186; EGF_2; FALSE_NEG.
KW EGF-like domain; Repeat; Transmembrane; Alternative splicing.
FT TRANSMEM 1 29
FT NON_TER 1
FT TRANSMEM 9 182
FT TRANSMEM 162 182
FT TRANSMEM 392 410
FT TRANSMEM 467 515
FT DOMAIN 467 515
FT DOMAIN 616 660
FT DOMAIN 664 716
FT DOMAIN 804 885
FT DOMAIN 1038 1153
FT DOMAIN 471 483
FT DISULFID 471 498
FT DISULFID 500 514
FT DISULFID 620 633
FT DISULFID 627 644
FT DISULFID 646 659
FT DISULFID 668 680
FT DISULFID 676 700
FT DISULFID 702 715
FT CARBOHYD 123 123
FT CARBOHYD 234 234
FT CARBOHYD 380 390
FT CARBOHYD 461 461
FT CARBOHYD 737 737
FT CARBOHYD 750 750
FT CARBOHYD 812 812
FT CARBOHYD 992 992
FT VARSPLIC 488 496
SQ SEQUENCE 1310 AA; 145670 MW; 4A2BA8E07E03EC66 CRC64;

Query Match 11.9%; Score 65.5; DB 1; Length 1310;
Best Local Similarity 20.8%; Pred. No. 63;

Matches 26; Conservative 20; Mismatches 38; Indels 41; Gaps 4;

OY 2 TMRGALLVLLVLTQALGVK-----MAETCP-----IFYDVFVAVANG 39
| | | | | : : : : : | | | | | : : : : :
DB 692 TLVNLMPGCGIEDYKLGVDGRSCQLITETCPGSDGSGSRELPMNQTLGEMFFQYNNH 751
| | | | | : : : : : | | | | | : : : : :
OY 40 NEILLDSLTKVNATEPERTAMKKIQQDCVY---GLISRVLDGLVMTTSSSKDCKGE 95
| | | | | : : : : : | | | | | : : : : :
DB 752 SK-----EVAAGGVLEKGFTRONNFARGLDQQLPDGLVAVTPLENCLTEE 796
| | | | | : : : : : | | | | | : : : : :
OY 96 AVQNT 100
| | | | | : : : : : | | | | | : : : : :
DB 797 ISEPT 801
| | | | | : : : : : | | | | | : : : : :

RESULT 15
NHAB_HAEPIN STANDARD; PRT; 522 AA.
ID NHAB_HAEPIN 544706;
AC P44706;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NA(+) / H(+) ANTIPORTER 2 (SODIUM/PROTON ANTIporter 2).
GN NHAB OR H10427.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Heddlom E., Cotton M.D.,
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghagen N.S.M.,

RESULT	2	
Q9TYW0		
ID	Q9TYW0	PRELIMINARY;
AC	Q9TYW0;	PRT; 1247 AA.
DT	01-MAY-2000	(TREMBLRel. 13, Created)
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)

```

DE 01-MAY-2000 (TREMBLrel. 13, last annotation update)
DE CLU-1.
CN CLU-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-N2;
RC Wickham J.O., Moulder G., Barstead R., Clarke M.;
RT "Cloning and disruption of the clu-1 gene of C. elegans."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF101078; AAC72406.1;
SQ SEQUENCE 1247 AA; 139992 MW; BC2591EF25D9FAF8 CRC64;

Query Match 14.1%; Score 78; DB 5; Length 1247;
Best Local Similarity 28.0%; Pred. No. 19;
Matches 26; Conservative 14; Mismatches 37; Indels 16; Gaps 2;

OY 28 IFYDFFAVANGNELLDLSITKYNATPEPTAMKKIDCYENGLI-----SRV 77
DB 384 IHADYNAARQVQSLDGNLAINGEDEKTHM-----YIMNIFPSLGEVDNDHYKE 437
OY 78 LDGLVMTTSSSKDCGEAVONTVEDIKNTLG 110
DB 438 LGDDAANAFAATSTDLQVAFATLDPKINTLG 470

RESULT 3
O26216 PRELIMINARY; PRT; 2771 AA.
ID O26216;
AC O26216;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE RHOPTRY PROTEIN.
OS Plasmodium yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5861;
RN [1]
RP SEQUENCE OF 379-2771 FROM N.A.
RA STRAIN-YM;
RC MEDLINE-97077455; PubMed-8920022;
RX Sinha K.A., Keen J.K., Ogun S.A., Holder A.A.;
RT "Comparison of two members of a multigene family coding for high-
RT molecular mass rhoptry proteins of Plasmodium yoelii."
RL Mol. Biochem. Parasitol. 76:329-332(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-YM;
RC Green J.L., Holder A.A.;
RT "Structure of the B8 gene encoding a high molecular mass rhoptry
RT protein of Plasmodium yoelii."
RL Mol. Biochem. Parasitol. 0:0-0(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-YM;
RC Holder A.A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U36927; AAB41263.3;
SQ SEQUENCE 2771 AA; 325640 MW; C0CCB9AB67ACF36 CRC64;

Query Match 13.9%; Score 76.5; DB 5; Length 2771;
Best Local Similarity 35.1%; Pred. No. 72;
Matches 27; Conservative 12; Mismatches 25; Indels 13; Gaps 4;

OY 34 FAVANGNELLDLSITKYNATPEPTAMKKIDCYE-NGLISRVLDLVTITSSSKD 91
DB 2102 FSENNNTLOSRRKIKELTNAFNAE---IKIKEDIKTEKNGINKL-----IETRKD 2150

```

```

OY 92 CMGEAVONTVEDIKLNT 108
DB 2151 CMLEFYTKLVETLKIRK 2167

RESULT 4
O9M7A8 PRELIMINARY; PRT; 945 AA.
ID O9M7A8;
AC O9M7A8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE LRR RECEPTOR-LIKE PROTEIN KINASE.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV, XANTH;
RC Cho H.S., Yoon G.M., Liu J.R., Pai Lee H.S.;
RT "Expression of NLRK1 encoding a LRR receptor-like kinase is
RT stimulated by abiotic and biotic stresses."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF142596; AAF66615.1;
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_tyr.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00560; LRR_6.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR_7.
DR SMART: SM00369; LRR_TYR_1.
DR SMART: SM00221; STRCG_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW transferase.
SQ SEQUENCE 945 AA; 101787 MW; 30741AFEB8909145E CRC64;

Query Match 13.6%; Score 75; DB 10; Length 945;
Best Local Similarity 26.1%; Pred. No. 28;
Matches 24; Conservative 18; Mismatches 38; Indels 12; Gaps 4;

OY 17 ALGVKMAEPPIYDVFVAVANGNELLDLSITKYNATPEPTAMKKIDCYENGLIS 76
DB 174 AVSANITGITPDPDFAFSLTN-----LHISFNNGCSLPSFSQSIQSLWL-NOLKOR 227
OY 77 VLDGLV-----MTTSSKDCGEAVONTVED 103
DB 228 -LNSIAVIONMTQLTRISGCANAFSPILPD 258

RESULT 5
O9A0D8 PRELIMINARY; PRT; 404 AA.
ID O9A0D8;
AC O9A0D8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PURATIVE THIAMINE BIOSYNTHESIS PROTEIN.
GN THII OR SPY0817.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

```

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370;
RA MEDLINE=21192684; PubMed=11296296;
RX Ferretti J.J., Moshen W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najaf F.Z., Ren O., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006532; AAK3753.1; -
KM Complete proteome.
SQ SEQUENCE 404 AA; 44730 MW; 9BC8B801385463AB CRC64;

Query Match 13.0%; Score 72; DB 2; Length 404;
Best Local Similarity 27.5%; Pred. No. 21;
Matches 14; Conservative 17; Mismatches 18; Indels 2; Gaps 1;

OY 54 TEPERTAMKIDDCYVENGISRVLDGLVMTTSSKDCMGAVONTVEDL 104
DB 355 TNPKLGNAKYECPDIDELVORAVSGIVTEI--TPELVNDEVNLDAL 403

RESULT 6
O9CHM8 PRELIMINARY; PRT; 800 AA.
ID O9CHM8;
AC O9CHM8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLYCOCEN PHOSPHORYLASE (EC 2.4.1.1) (ALPHA-D-GLUCAN PHOSPHORYLASE)
DE (STARCH PHOSPHORYLASE).
GN GLCP.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA Bolotin A., Wincker P., Mauer S., Jallion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus lactis*.";
RL Genome Res. 0:0-0(2001).
CC -1- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES. HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL PROPERTIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: (1,4-ALPHA-D-GLUCOSYL)(N) + ORTHOPHOSPHATE = (1,4-ALPHA-D-GLUCOSYL)(N-1) + ALPHA-D-GLUCOSE 1-PHOSPHATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: TO THE GLYCOCEN PHOSPHORYLASE FAMILY.
DR EMBL: AE006303; AAK04798.1; -
DR InterPro: IPR000811; Phosphorylase.
DR Pfam: PF00343; phosphorylase; 1.
DR PROSITE: PS00102; PHOSPHORYLASE; 1.
KM Carbohydrate metabolism; Complete proteome; Glycosyltransferase; Pyridoxal phosphate; Transferase.
SQ SEQUENCE 800 AA; 91812 MW; A9A3CC9A81D0AD27 CRC64;

Query Match 13.0%; Score 72; DB 2; Length 800;
Best Local Similarity 28.3%; Pred. No. 48;
Matches 17; Conservative 14; Mismatches 27; Indels 2; Gaps 1;

OY 31 DVEFAANNGNELLDLSTKRVATEPERTAMKIDDCYVENGISRVLDGLVMTTSSSK 90
DB 665 ELFEAGDGNFV--EGLTKDEVYEVYRNGNYNARDIYQNPVNRILNALIDGIVPNIK 722

RESULT 7
O9ES84 PRELIMINARY; PRT; 119 AA.
ID O9ES84;
AC O9ES84;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SNRP CORE PROTEIN SMX5D (FRAGMENT).
GN SMX5D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.D2/SGSNU; TISSUE=TESTIS;
RA Zhang J., Ma Z.R., Meeker N.D., Teuscher C.;
RT "Identification and characterization of alternatively spliced Smx5 isoforms in mouse.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF204157; AAG31434.1; -
DR InterPro: IPR001163; snRNP_Sm.
DR Pfam: PF01423; Sm; 1.
FT NON_TER 1 1
SQ SEQUENCE 119 AA; 13228 MW; B326E5EC16D6F204 CRC64;

Query Match 12.9%; Score 71; DB 11; Length 119;
Best Local Similarity 23.2%; Pred. No. 6.2;
Matches 22; Conservative 21; Mismatches 28; Indels 24; Gaps 4;

OY 6 ALLVIALTYQALGYVMAETCP---IFYDVFVAVANGNEL----- 43
DB 1 ALVSVAMGSGQAPGFSVQPCPSLFYS-FPKSVGKVVELKNDLSICGTLSVDQY 59
OY 44 LDLSLTKRVATEPERTAMKIDDCYVENGISRV 77
DB 60 LNKILDISVTDEPKYPHLSVKNCFIRGSVRYV 94

RESULT 8
O55959 PRELIMINARY; PRT; 1252 AA.
ID O55959;
AC O55959;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 136.1 KDA PROTEIN.
GN SLR0697.
OS *Synechocystis* sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N., Sugitani M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nanno K., Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";

RL DNA Res. 3:109-136(1996).
 EMBL: D64005; BAA10729.1.
 DR InterPro: IPR002821; Hydantoinase_A.
 DR InterPro: IPR003692; Hydantoinase_B.
 DR Pfam: PF01968; Hydantoinase_A; 1.
 DR Pfam: PF02538; Hydantoinase_B; 1.
 KW Hypothetical protein; Complete proteome.
 SW SEQUENCE 1252 AA; 136051 MW; 07ECCA495CDB25F CRC64;

Query Match 12.98; Score 71; DB 2; Length 1252;
 Best Local Similarity 30.48; Pred. No. 1e+02;
 Matches 24; Conservative 13; Mismatches 38; Indels 4; Gaps 2;

OY 21 KMAETCPFYDFVFAVANGNELLDLSTKYNATEPERTARKIDOCYENGCLSRVLDG 80
 DB 49 KILSENPELYD-DVAIHGRITLGLSINERIA--PEKSVKMGTTVATNALLEQKDDP 104
 OY 81 LVMTTSSKDCGEGAVON 99
 DB 105 LVLMTOGFRDGLAIAYON 123

RESULT 9
 OYX242 PRELIMINARY; PRT; 295 AA.
 AC OYX242
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 14, Last sequence update)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN TM1717.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID-2336;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-MS88 / DSM 3109;
 RA MEDLINE-9287316; Pubmed-10360571.
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 McDonald L., Utterback T.R., Malek J.A., Liner K.D., Garrett M.M.,
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.W., White O.,
 RT Evidence for lateral gene transfer between Archaea and Bacteria from
 RL genome sequence of Thermotoga maritima.
 DR EMBL: A001333-329(1999)
 DR EMBL: A001331; A0036783.1;
 DN TIGR: TM1717; Gene.
 KW Complete proteome.
 SW SEQUENCE 295 AA; 33641 MW; E8A9E7396C85848 CRC64;

Query Match 12.88; Score 70.5; DB 2; Length 295;
 Best Local Similarity 26.68; Pred. No. 21;
 Matches 25; Conservative 19; Mismatches 31; Indels 19; Gaps 4;

OY 12 LVYQALGVKMAETCPFYDFVFAVANGNELLDLSTKYNATEPERTARKIDOCYEN 71
 DB 84 LVYT---VKNPESTYTIIDKFLVLEKKELELVAVVANKKMDLYDED--DLKVRD----- 132
 OY 72 GLISRYLDGLVMTTSSKDCGEGAVONTVEDK 105
 DB 133 ---LEETISGLIPYIKTSIKTKGMS-----ISELK 158

RESULT 10
 ID O48678 PRELIMINARY; PRT; 388 AA.
 AC O48678
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE F316.4 PROTEIN.
 GN F316.4.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID-3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Federspiel N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Wu M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
 RA Oji O., Osborne B.I., Shinn P., Sun H., Tortum M., Vysotskaya V.S.,
 RA Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002396; AAC00573.1;
 DR HSSP: P25685; 1HDJ.
 DR InterPro: IPR001623; DnaJ_N.
 DR Pfam: PF00226; DnaJ_1.
 DR SMART: SM00271; DnaJ_1.
 DR PROSITE: PS00636; DnaJ_1; UNKNOWN_1.
 DR PROSITE: PS50076; DnaJ_2; 1.
 SW SEQUENCE 388 AA; 44023 MW; DD815C3AD9BDAF01 CRC64;

Query Match 12.88; Score 70.5; DB 10; Length 388;
 Best Local Similarity 29.68; Pred. No. 29;
 Matches 24; Conservative 12; Mismatches 30; Indels 15; Gaps 3;

OY 1 DTMGALLVLLVYQALGVKMAETCPFYDFVFAVANGNELLDLSTKRV 51
 DB 132 EALNGRVSDPLVIGVNSKVEKCAHFAVY---TISEELLYFDDEANSGLSLQWA 188
 OY 52 NATPERTARKKI--QDCYV 69
 DB 189 QAKDPETAFPKLIDGFOQCEV 209

RESULT 11
 OYXBXO PRELIMINARY; PRT; 245 AA.
 AC OYXBXO
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE R105E-5-PHOSPHATE ISOMERASE A.
 GN XPF2015.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID-2371;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RA MEDLINE-20365717; Pubmed-10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin L.E.A., Carreiro D.M., Carier H.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carro C.M.,
 RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.U., Cristofani M., Dias-Neto E., Docena C., El-Dorty H.,
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferraz J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Honneisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miranca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silvestri W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*";
 RL Nature 406:151-159(2000).
 DR EMBL; AE004020; AAF84817.1; -;
 KW Complete proteome.
 SQ SEQUENCE 245 AA; 26561 MW; 2D2079AD223900E5 CRC64;

Query Match 12.7%; Score 70; DB 2; Length 245;
 Best Local Similarity 27.0%; Pred. No. 19;
 Matches 27; Conservative 13; Mismatches 22; Indels 38; Gaps 4;

OY 47 SLTKVATP-----ERTAMKTIQDCYVENGSLISRV----- 77
 DB 5 TYTRVNSTPANAACAPSHPIIINPMSEAKRRAEKALE--YVENDMTIGVGTSTVAV 62
 OY 78 -LDGLVMT-----TSSSKDCGGAQVONTVEDIKINTIG 110
 DB 63 FIDALGRTPKRIKGAVSSEOSTAHLKHGIEVLELHNHG 102

RESULT 12

ID 09HI70 PRELIMINARY; PRT; 770 AA.
 AC 09HI70;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE RIBONUCLEOTIDE REDUCTASE.
 GN TAI475.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
 UC Thermoplasma;
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger *Thermoplasma*
acidophilum.";
 RL Nature 407:508-513(2000).
 DR EMBL; AA445067; CAC12593.1; -;
 DR InterPro; IPR000788; Ribonucleo_red.
 DR Pfam; PF00317; ribonucleo_red; 3.
 DR PRINTS; PRO1183; RIBORDTASEM1.
 KW Complete proteome.
 SQ SEQUENCE 770 AA; 86925 MW; 4A5E1E1945842CF4 CRC64;

Query Match 12.5%; Score 69; DB 1; Length 770;
 Best Local Similarity 27.4%; Pred. No. 93;
 Matches 26; Conservative 17; Mismatches 26; Indels 26; Gaps 5;

OY 30 YDFFAVANGN---ELLDSLTKVNATEPERTAMKTIQDCYVENGSLISRVLDGLVMTT 85
 DB 27 YKAMLSVKNGTMDADQADKVVARK--DKRPSVEEID--VVEDVMTSKIDGKTFTD 83

OY 86 ISSS-----KDCGGAQVONTVEDIKL 106
 DB 84 VAKSYILYREKRAIREKELMG-----VKDDLKL 113

RESULT 13

P74911
 ID P74911 PRELIMINARY; PRT; 857 AA.
 AC P74911;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE RIBONUCLEOTIDE REDUCTASE.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
 UC Thermoplasma;
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tauer A., Benner S.A.;
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(0).
 DR EMBL; U73619; AAB18239.1; -;
 DR InterPro; IPR000788; Ribonucleo_red.
 DR Pfam; PF00317; ribonucleo_red; 3.
 DR PRINTS; PRO1183; RIBORDTASEM1.
 SQ SEQUENCE 857 AA; 96960 MW; 914406F575B2B93 CRC64;

Query Match 12.5%; Score 69; DB 1; Length 857;
 Best Local Similarity 27.4%; Pred. No. 11e+02;
 Matches 26; Conservative 17; Mismatches 26; Indels 26; Gaps 5;

OY 30 YDFFAVANGN---ELLDSLTKVNATEPERTAMKTIQDCYVENGSLISRVLDGLVMTT 85
 DB 25 YKAMLSVKNGTMDADQADKVVARK--DKRPSVEEID--VVEDVMTSKIDGKTFTD 81
 OY 86 ISSS-----KDCGGAQVONTVEDIKL 106
 DB 82 VAKSYILYREKRAIREKELMG-----VKDDLKL 111

RESULT 14

ID 09YAUI PRELIMINARY; PRT; 895 AA.
 AC 09YAUI;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 895AA LONG HYPOTHETICAL DNA-DIRECTED RNA POLYMERASE SUBUNIT A'.
 GN APE1853.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 UC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, *Aeropyrum pernix* K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000062; BAA80857.1; -;
 DR InterPro; IPR000722; RNA_pol_A.
 DR Pfam; PF00623; RNA_pol_A; 1.
 KW DNA-directed RNA polymerase; Complete proteome.
 SQ SEQUENCE 895 AA; 101740 MW; 1E01BC064BD0AAE CRC64;

Query Match 12.5%; Score 69; DB 1; Length 895;
 Best Local Similarity 30.9%; Pred. No. 1.1e+02;
 Matches 21; Conservative 12; Mismatches 21; Indels 14; Gaps 2;

OY 31 DVEFAVANGNELLDLSLTKVNATEPERTAMKTIQDCYVENGSLISRVLDGLVMTTSSSK 90

Wed May 8 09:01:34 2002

us-09-662-784-6.rspt

Page 6

```
Db      806 EVFFHAAGREGLV-----TAVKTSQSGYMORRLNALDP-IVVYDGTVR 851
QY      91 DCMGEAVQ 98
Db      852 DLYGNLIQ 859
```

RESULT 15

```
Q92519      PRELIMINARY:      PRT:      343 AA.
ID Q92519;
AC Q92519;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GS3955 (GS3955 PROTEIN).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CANCELLOUS BONE;
RA Ohno I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: D87119; BAA13250.1;
DR EMBL: BC002637; AAH02637.1;
DR HSSP: P08631; IAD5.
DR InterPro: IPR000719; Euk_Pkinase.
DR Pfam: PF00069; pkinase; 2.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase;
SQ SEQUENCE 343 AA; 38800 MW; EF8B7365DACB84FA CRC64;
```

Query Match 12.48; Score 68.5; DB 4; Length 343;
Best Local Similarity 35.28; Pred. No. 40;
Matches 19; Conservative 9; Mismatches 21; Indels 5; Gaps 2;

```
OY      24 ETCPIFYDVFAVA--NGNELLDISTKYNATEPERTAMK--KIQDCYENG 72
Db      151 EARLFYQIASAVAHCHDGGLVDRDKLKEFKDEERTVKNLESLEDAITLIG 204
```

Search completed: May 7, 2002, 12:16:47
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:14:13 ; Search time 23.78 Seconds
(without alignments)
345.758 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 111

Sequence: 1 DTRGALVLLALVLTQALGV.....CMGEAVQNTVEDIKLNTLGR 111

Scoring table: OLIGO

Gapop 60.0 , Gapept 60.0

Searched: 522463 seqs, 74073290 residues

Word size : 0

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.GeneSeq_1101.*
1: /SID8/gcgdata/geneSeq/geneSeq/AA1980.DAT.*
2: /SID8/gcgdata/geneSeq/geneSeq/AA1981.DAT.*
3: /SID8/gcgdata/geneSeq/geneSeq/AA1982.DAT.*
4: /SID8/gcgdata/geneSeq/geneSeq/AA1983.DAT.*
5: /SID8/gcgdata/geneSeq/geneSeq/AA1984.DAT.*
6: /SID8/gcgdata/geneSeq/geneSeq/AA1985.DAT.*
7: /SID8/gcgdata/geneSeq/geneSeq/AA1986.DAT.*
8: /SID8/gcgdata/geneSeq/geneSeq/AA1987.DAT.*
9: /SID8/gcgdata/geneSeq/geneSeq/AA1988.DAT.*
10: /SID8/gcgdata/geneSeq/geneSeq/AA1989.DAT.*
11: /SID8/gcgdata/geneSeq/geneSeq/AA1990.DAT.*
12: /SID8/gcgdata/geneSeq/geneSeq/AA1991.DAT.*
13: /SID8/gcgdata/geneSeq/geneSeq/AA1992.DAT.*
14: /SID8/gcgdata/geneSeq/geneSeq/AA1993.DAT.*
15: /SID8/gcgdata/geneSeq/geneSeq/AA1994.DAT.*
16: /SID8/gcgdata/geneSeq/geneSeq/AA1995.DAT.*
17: /SID8/gcgdata/geneSeq/geneSeq/AA1996.DAT.*
18: /SID8/gcgdata/geneSeq/geneSeq/AA1997.DAT.*
19: /SID8/gcgdata/geneSeq/geneSeq/AA1998.DAT.*
20: /SID8/gcgdata/geneSeq/geneSeq/AA1999.DAT.*
21: /SID8/gcgdata/geneSeq/geneSeq/AA2000.DAT.*
22: /SID8/gcgdata/geneSeq/geneSeq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	111	13	TRFP Chain #2 with
2	111	100.0	111	21	T cell reactive fe
3	111	100.0	111	21	Feline human TRFP
4	111	100.0	111	21	Cat TRFP chain 2 1
5	111	100.0	111	21	Human TRFP chain 2
6	109	98.2	109	14	TRFP chain 2 (with
7	109	98.2	109	14	Human T cell react
8	109	98.2	109	20	Felis sp. allergen
9	99	89.2	111	12	TRFP chain 2 - lon
10	92	82.9	92	21	T cell reactive fe
11	92	82.9	92	21	Feline human TRFP

12	92	82.9	92	21	AAV51475	Human TRFP chain 2
13	83	74.8	97	12	AAAR2123	TRFP chain 2 - tru
14	83	74.8	109	21	AAAB28934	T cell reactive fe
15	83	74.8	109	21	AAAB7674	Feline human TRFP
16	83	74.8	109	21	AAV90104	Cat TRFP chain 2 s
17	83	74.8	109	21	AAV51471	Human TRFP chain 2
18	83	74.8	110	13	AAAR27370	TRFP Chain #2 with
19	81	73.0	97	21	AAAB28935	T cell reactive fe
20	81	73.0	97	21	AAAB7675	Feline human TRFP
21	81	73.0	97	21	AAV90105	Cat TRFP chain 2 t
22	81	73.0	97	21	AAV51472	Human TRFP chain 2
23	79	71.2	109	12	AAAR2122	TRFP I chain 2 - s
24	66	59.5	92	18	AAAR27382	Chain 2 of major c
25	66	59.5	92	20	AAV25532	Human MHC class II
26	64	57.7	90	21	AAAB28938	T cell reactive fe
27	64	57.7	90	21	AAAB7679	Feline human TRFP
28	64	57.7	90	21	AAV51476	Human TRFP chain 2
29	51	45.9	101	13	AAAR27372	TRFP Chain #2 C2ST
30	37	33.3	51	21	AAAB28939	T cell reactive fe
31	37	33.3	51	21	AAAB7680	Feline human TRFP
32	37	33.3	51	21	AAV51477	Human TRFP chain 2
33	37	33.3	82	13	AAAR27371	TRFP Chain #2 C2ST
34	32	28.8	82	21	AAAB28940	T cell reactive fe
35	32	28.8	82	21	AAAB7681	Feline human TRFP
36	32	28.8	82	21	AAV90106	Cat TRFP chain 2 g
37	32	28.8	82	21	AAV51478	Human TRFP chain 2
38	26	23.4	26	14	AAAB3544	Peptide Z. Felis.
39	26	23.4	26	14	AAAR41977	Human T cell react
40	26	23.4	26	21	AAAB28943	Peptide Z derived
41	26	23.4	26	21	AAAB7684	Feline human TRFP
42	26	23.4	26	21	AAV90109	Cat TRFP derived p
43	26	23.4	26	21	AAV51481	Human TRFP derived
44	26	23.4	96	14	AAAB3548	Recombinant YZX.
45	26	23.4	96	21	AAAB28979	Peptide YZX. Fel

ALIGNMENTS

RESULT 1	
AAAR27369	standard: protein; 111 AA.
XX	
AC	AAAR27369;
XX	
DT	20-MAY-1998 (first entry)
XX	
DE	TRFP Chain #2 with C2 leader sequence.
XX	
KW	T cell reactive feline protein; cat allergy; allergc; IGF;
XX	
OS	Felis domesticus.
XX	
FH	Key
FT	peptide
FT	1.19
FT	/label= C2 leader
FT	19.111
FT	/label= TRFP chain #2
PN	
PD	WO9215613-A.
XX	
PD	17-SEP-1992.
XX	
PF	20-FEB-1992; 92WO-US01344.
XX	
PR	28-FEB-1991; 91US-0662193.
XX	
PA	(IMMU-) IMMUNOLOGIC PHARM CORP.
XX	
PI	Bond J, Kuo M;
XX	
DR	WPI; 1992-331670/40.

XX Modified human T-cell reactive feline protein - stimulates T-cell
PT In individuals allergic to cats and shows reduced
PR histamine-releasing properties
PS
XX Claim 1; Fig 1; 35pp; English.
XX This sequence represents a modified human T-cell reactive feline
CC protein which stimulates T-cells from an individual who is allergic
CC to cats, but which interacts with human IgE to a lesser extent than
CC does affinity purified TRFP. The protein is modified by treating
CC with either a mild alkali (pH 12.5-13.5, KOH, NaOH, LiOH or tertiary
CC amines) or an enzyme which removes O-linked groups (carbohydrate
CC moieties). It is useful in desensitizing people who are allergic to cats.
SQ Sequence 111 AA;
Query Match 100.0%; Score 111; DB 13; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.5e-100;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DTWKGALLVLAIVTQALGVKMAETCPFYDVFFAVANGNELLDLSTFKVNATEPERTA 60
DB 1 dtmrgallviallvltqalgykmaetcpifdyvffavangnellldstltnaeperta 60
OY 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCMGAEAVONTVEDLKTNTLGR 111
DB 61 mkkldcyvengllsrvidglvmtlssskdcmgeavqntvedlknltlgr 111
RESULT 2
ID AAB28933 standard; Protein; 111 AA.
XX
AC AAB28933;
XX
DT 29-JAN-2001 (first entry)
XX
DE T cell reactive feline protein chain 2 long form.
XX
KW Cat; allergy; human T cell reactive feline protein; hTRFP;
XX immunotherapy.
XX
OS Fells sp.
XX
PN US6120769-A.
XX
PD 19-SEP-2000.
XX
PE 28-APR-1995; 9505-0431184.
XX
PR 02-SEP-1994; 9405-0300928.
XX
PR 03-NOV-1989; 8905-0431565.
XX
PR 28-FEB-1991; 9105-0662276.
XX
PR 13-DEC-1991; 9105-0807529.
XX
PR 25-MAR-1992; 9205-0857311.
XX
PR 15-MAY-1992; 9205-0884718.
XX
PR 15-JAN-1993; 9305-0006116.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Gelfer ML, Garman RD, Greenstein JL, Bond JF.
XX
DR WPI: 2000-601477/57.
XX
DR N-PSDB; AAC60103.
XX
PT Detecting, preventing and treating sensitivity to cat protein allergen
PT comprises combining a biological sample with a human T cell reactive
XX feline protein and determining the extent of binding that occurs -
XX
PS Claim 1; Figure 3; 106pp; English.

CC The present invention relates to the detection of sensitivity to a cat
CC protein allergen by combining a blood sample from a subject with a
CC peptide of human T cell reactive feline protein (hTRFP). This method
CC and the hTRFP peptides are useful for diagnosing, preventing and
CC treating cat allergies by reducing or abolishing an individual's
CC allergic response to a cat allergen. DNA encoding the TRFP may be
CC used as probes to locate equivalent sequences present in other species.
CC These may further be used to study the mechanism of immunotherapy of
CC cat allergy, and to design modified derivatives, analogues or
CC functional equivalents useful in immunotherapy. The present
SQ sequence was used in the invention.
SQ Sequence 111 AA;
Query Match 100.0%; Score 111; DB 21; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.5e-100;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DTWKGALLVLAIVTQALGVKMAETCPFYDVFFAVANGNELLDLSTFKVNATEPERTA 60
DB 1 dtmrgallviallvltqalgykmaetcpifdyvffavangnellldstltnaeperta 60
OY 61 MKRIQDCYVENGILSRVLDGLVMTTSSSKDCMGAEAVONTVEDLKTNTLGR 111
DB 61 mkkldcyvengllsrvidglvmtlssskdcmgeavqntvedlknltlgr 111
RESULT 3
ID AAY87673 standard; Protein; 111 AA.
XX
AC AAY87673;
XX
DT 22-AUG-2000 (first entry)
XX
DE Feline human TRFP chain 2 long form protein.
XX
KW T-cell reactive feline protein; TRFP; Fel d I; cat allergen;
XX antiallergic; T cell stimulator; diagnostic; immunotherapy.
XX
OS Fells sp.
XX
PN US6048962-A.
XX
PD 11-APR-2000.
XX
PE 27-APR-1995; 9505-0430014.
XX
PR 02-SEP-1994; 9405-0300928.
XX
PR 03-NOV-1989; 8905-0431565.
XX
PR 28-FEB-1991; 9105-0662276.
XX
PR 13-DEC-1991; 9105-0807529.
XX
PR 25-MAR-1992; 9205-0857311.
XX
PR 15-MAY-1992; 9205-0884718.
XX
PR 15-JAN-1993; 9305-0006116.
XX
PA (IMMU-) IMMULOGIC PHARM CORP.
XX
PI Kuo M, Rogers BU, Gelfer ML, Morgenstern JP, Brauer AM,
XX Greenstein JL, Griffith JF, Garman RD;
XX
DR WPI: 2000-316905/27.
XX
DR N-PSDB; AAA12244.
XX
PT New human T cell reactive feline protein useful for reducing or
PT abolishing individual's allergic response to cat allergen comprising
XX two different covalently linked chains -
XX
PS Claim 2; Column 75-76; 106pp; English.
XX This invention describes a novel naturally occurring cat protein allergen
CC (I), human T cell reactive feline protein (TRFP), comprising two

different covalently linked peptide chains with a molecular weight of 20 kD, 40 kD or 110 kD under non-reducing conditions and 5 kD or 10-18 kD under reducing conditions. The products of the invention have antiallergic activity and act as human T cell stimulants. TRFP is useful for reducing or preventing the adverse effects of cat allergens on cat allergic individuals and in ex vivo diagnostic tests to determine which peptides cause sensitivity so as to selectively use them to desensitize a cat sensitive individual. Purified TRFP is also useful for studying the mechanism of immunotherapy of cat allergy and to design modified derivatives, analogs or functional equivalents that are more useful in immunotherapy against cat allergy. DNA sequences encoding TRFP are useful as probes to locate equivalent sequences present in other species (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics and/or therapeutics. Fully defined and characterized TRFP provides complete and a very simple desensitization therapy. This sequence represents a human T cell reactive feline protein (also known as Fel d 1) chain 2, long form which is described in the method of the invention.

Query Match	100.0%;	Score 111;	DB 21;	Length 111;
Best Local Similarity	100.0%;	Pred. No. 2.5e-100;		
Matches 111;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

RESULT	4
AAV90103	
ID	AAV90103 standard; Protein; 111 AA.

AC	AA90103;
XX	
DT	13-JUL-2000 (first entry)

DE Cat TRFP chain 2 long form protein sequence.

KM Cat; TRF; human T-cell reactive feline protein; cat protein allergen;
KM house dust; Fel d I; cat allergy; Felis domesticus sensitivity; therapy;
KM diagnosis; goat; sheep; horse; rabbit; dog.

OS *Felis domesticus*.

FH	Key	Location/qualifiers
FT	Peptide	1..19
FT		/note= "signal peptide"
FT	Protein	20..111
FT		/note= "mature TRFP chain 2 long form"

PN US6025162-A.

PD 15-FEB-2000

PF 28-APR-1995;

02-SEP-1994;

PR 28-FEB-1991;
 PR 13-DEC-1991;

PR 23-MAR-1992
PR 15-MAY-1992

LC-CHAN-1993/

XX
XX

104

unvergleichlich ist

XX WPI; 2000-181812/16.
DR N-PSDB; AAA07437.
DR

PT New human T cell reactive feline protein, useful for desensitizing cat
PT allergic individuals to cat allergens -

PS Claim 1; Fig. 3; 108pp; English.

CC This sequence is a peptide chain of the human feline reactive
CC protein (TRFP) of the invention. The protein is a cat protein allergen,
CC and was isolated from a vacuum bag extract obtained by affinity
CC purification of house dust collected from several homes with cats. TRFP
CC is composed of two covalently linked peptide chains, and is also referred
CC to as Fel d 1. TRFP and its peptides are useful for reducing or
CC preventing the adverse effects that exposure to cat allergens normally
CC has on cat allergic individuals (i.e. to desensitise individuals to cat
CC allergens or block the effect of the allergens). TRFP is also used in
CC methods of diagnosing sensitivity to feline domesticus in an individual.
CC DNA sequences encoding TRFP can be used as probes to locate equivalent
CC sequences present in other species, e.g. goat, sheep, horse, rabbit and
CC dog, that may be useful in a diagnostic and/or therapeutic applications.
XX
XX Sequence 111 AA;

Query Match	100.0%	Score 111:	DB 21:	Length 111:	.
Best Local Similarity	100.0%	Pred. No. 2.5e-100:			
Matches 111: Conservative	0:	Mismatches	0:	Indels	0:
				Gaps	0:

Qy 1 DTMGALIVALLVTPALCVKNAECPIFEYDVFFVANGNLLDLSLTVTNATPEERIA 60
Db 1 dltmgallviallvtagialgymaetcpifdyffvfangnellldlsltkvnatpeerta 60
Qy 61 MKKIQDCYENGILSRVLDGLVYTTITSSKDKDGEAVONTVEDIKLTLICR 111
Db 61 mkkigdcyengilsvrlglwmtitsskdkdmeaavontvediklntlcr 111

RESULT	5
AAV51470	
ID	AAV51470 standard; Protein; 111 AA

AC AAY51470;

DT 22-MAY-2000 (first entry)

Human TRFP chain 2 (long form) protein fragment

KW T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;

KW sensitivity; cat protein allergen; human; chain 2
 XY

OS Homo sapiens.
XY

US60199/2-A.
PN
XX

PD 01-FEB-2000
XX

02 SEP 1994
XX

PR 28-FEB-1991

PR 25-MAR-1992;

PR 15-JAN-1993

PA (IMMU-) IMM

PI Garman RD,

DR WPI; 2000-14

Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gelfer ML; WPI; 2000-146862/13.

WPI; 2000-146862/13.

DR N-PSDB; AAZ88617.
 XX Peptides of human T cell reactive feline protein for treating
 PT sensitivity to cat protein allergens comprise at least one T cell
 PT epitope recognized by a T cell receptor specific for the human T cell
 XX reactive feline protein -
 XX
 PS Claim 1; Column 75-78; 105pp; English.
 CC This invention describes a novel peptide (1) of human T cell reactive
 CC feline protein (TRFP) having at least one T cell epitope recognized
 CC by a T cell receptor specific for the human T cell reactive feline
 CC protein. The peptide consisting of at least 7-30 amino acids, and having
 CC an amino acid sequence derived from an amino acid sequence comprising 94,
 CC 96, 97, 109, or 111 residues given in the specification. The peptides
 CC down regulate the immune response to the allergen. The peptides have
 CC reduced immunoglobulin E binding and induce T cell responsiveness. The
 CC peptide (1) is useful in compositions for creating sensitivity to a cat
 CC protein allergen in a subject. This sequence represents the human TRFP
 CC chain 2 (long form).
 CC
 SQ Sequence 111 AA;
 Query Match 100.0%; Score 111; DB 21; Length 111;
 Best Local Similarity 100.0%; Pred. No. 2.5e-100; Mismatches 0; Gaps 0;
 Matches 111; Conservative 0; Indels 0; Gaps 0;
 QY 1 DTRGALVLTALVLTQALGVKMAETCPIDYDFVAVANGNELLDLSTKRVNATPERTA 60
 DB 1 dtrgallvltallvtqalgvkmaetcpifdydfvavangnellldstlkrvatpepta 60
 QY 61 MKRIQCYENGILSRVLDGLVMTTSSSKDCMGAVONTVEDKINTLGR 111
 DB 61 mkridgyengilsrvidglvmttssskdcmgeavontvedkintlgr 111
 RESULT 6
 AAR36541
 ID AAR36541 standard; Protein; 109 AA.
 AC AAR36541;
 XX 12-AUG-1993 (first entry)
 DT TRFP chain 2 (with leader).
 DE
 XX Human T cell reactive feline protein; TRFP; leader A; leader B;
 KW epitope.
 XX
 OS Felis.
 FH
 FT Key Location/Qualifiers
 FT Peptide 1..17
 FT /label= leader_peptide
 PN W09308280-A.
 PD 29-APR-1993.
 XX 16-OCT-1992; 92WO-US08694.
 PF 16-OCT-1991; 91US-0777859.
 PR 13-DEC-1991; 91US-0807529.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Kuo M, Morgenstern JP, Morville M;
 PI Rogers BL;
 XX WPI: 1993-152473/18.
 DR N-PSDB; AAQ41558.
 XX

PT Recombitope peptide having T-cell stimulating activity - for the
 PT diagnosis and treatment of sensitivity to protein allergens;
 PT auto:antigens and protein antigens
 XX
 PS Disclosure; Fig 2; 73pp; English.
 CC Chains 1 and 2 of the TRFP have been recombinantly expressed in E.
 CC coli and purified. T cell epitope studies using overlapping peptide
 CC regions derived from the TRFP amino acids sequence were used to
 CC identify multiple T cell epitopes in each chain of TRFP.
 CC
 SQ Sequence 109 AA;
 Query Match 98.2%; Score 109; DB 14; Length 109;
 Best Local Similarity 100.0%; Pred. No. 2.1e-98; Mismatches 0; Gaps 0;
 Matches 109; Conservative 0; Indels 0; Gaps 0;
 QY 3 MRGALLVLTALVLTQALGVKMAETCPIDYDFVAVANGNELLDLSTKRVNATPERTAMK 62
 DB 1 mrgallvltallvtqalgvkmaetcpifdydfvavangnellldstlkrvatpeptamk 60
 QY 63 KIQDCYENGILSRVLDGLVMTTSSSKDCMGAVONTVEDKINTLGR 111
 DB 61 kridgyengilsrvidglvmttssskdcmgeavontvedkintlgr 109
 RESULT 7
 AAR41985
 ID AAR41985 standard; Protein; 109 AA.
 AC AAR41985;
 XX 21-APR-1994 (first entry)
 DT Human T cell reactive feline protein chain 2.
 DE
 XX Human: T cell; reactive; feline; protein; immune response; antigen;
 KW tolerance; mammal; Dermatophagoides; Felis; Ambrosia; Lolium; Canis;
 KW Cryptomeria; Alternaria; Alder; Betula; Quercus; Olea; Artemesia;
 KW Plantago; Parietaria; Blatella; Apis; Periplaneta; autoantigen.
 XX
 OS Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT Peptide 1..17
 FT /note= "Signal peptide"
 FT Protein 18..109
 FT /note= "Mature protein"
 PN W09319178-A.
 PD 30-SEP-1993.
 XX 25-MAR-1993; 93WO-US02462.
 PF 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 ER 15-JAN-1993; 93US-0006116.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Briner TJ, Garman RD, Geffter ML, Greenstein JL, Kuo M,
 PI Morville M;
 XX WPI: 1993-320744/40.
 DR N-PSDB; AAQ49535.
 XX New peptide(s) for inducing tolerance - comprise one or more
 PT epitope(s) of an allergen administered subcutaneously, for
 PT treating sensitivity to cats, bees, etc.
 XX
 PS Disclosure; Fig 2; 107pp; English.

XX This sequence represents chain 2 of human T cell reactive feline
CC protein (TRFP). Peptides derived from TRFP may be used in a
CC therapeutic composition which is useful in treating diseases which
CC involve an immune response to a protein antigen. This composition
CC may be used to induce tolerance in a mammal to Dermatophagoides,
CC Fells, Ambrosia, Lolium, Cyrtometaria, Alternaria, Alder, Betula,
CC Quercus, Olea, Artemisia, Plantago, Parietaria, Canis, Blatella,
CC Apis, Periplaneta and to autoantigens in humans.

XX
SQ Sequence 109 AA:

Query Match 98.2%; Score 109; DB 14; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.1e-98;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MRCALLVLLVLTQALGVKMAETCPFYDVFVAVANGNELLDLSLTKVATPEPTAMK 62
DB 1 mrgallvllalivcgalgvkmaetcpfydvffavangnellldsltkvatpeptamk 60
OY 63 KIDDCYVENGLSRVLGDMVTITSSKDCMGAVONTVEDLKLNTLGR 111
DB 61 kidqcyvenglsrvldgltvmtltsskdcmgavontvedlklntlgr 109

RESULT 8

AA25593
ID AAY25593 standard; protein; 109 AA.

AC AAY25593;
XX
DT 30-SEP-1999 (first entry)

XX Fells sp. allergen 1082946 Fel di chain 2 precursor protein fragment.

XX Major histocompatibility complex; class II; desensitizing; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mite; gerbil; vaccine; treatment; prevention; hypersensitivity.

XX Fells sp.

XX MO9934826-A1.

XX 15-JUL-1999.

XX 11-JAN-1999; 99WO-GB00080.

XX 21-SEP-1998; 98GB-0020474.

XX 09-JAN-1998; 98GB-0000445.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Kay AB, Larche M;

XX WPI; 1999-458255/38.

XX Desensitizing patients to polypeptide allergens

XX Example 6; Page 53; 117pp; English.

XX This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit

CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents the cat allergen 1082946 Fel di chain 2 precursor.

XX
SQ Sequence 109 AA:

Query Match 98.2%; Score 109; DB 20; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.1e-98;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MRCALLVLLVLTQALGVKMAETCPFYDVFVAVANGNELLDLSLTKVATPEPTAMK 62
DB 1 mrgallvllalivcgalgvkmaetcpfydvffavangnellldsltkvatpeptamk 60
OY 63 KIDDCYVENGLSRVLGDMVTITSSKDCMGAVONTVEDLKLNTLGR 111
DB 61 kidqcyvenglsrvldgltvmtltsskdcmgavontvedlklntlgr 109

RESULT 9

AA2121
ID AAR12121 standard; protein; 111 AA.

AC AAR12121;

XX 26-JUL-1991 (first entry)

XX TRFP chain 2 - long form.

XX Human T cell reactive feline protein; cat allergens.

XX Fells catus.

XX Key Location/Qualifiers

XX Peptide 3..19 /label= Leader sequence

XX Protein 20..111 /label= TRFP I long form

XX MO9106571-A.

XX 16-MAY-1991.

XX 02-NOV-1990; 90WO-US06548.

XX 03-NOV-1989; 89US-0431565.

XX (IMMU-) IMMUNOLOGIC PHARM COR.

XX Geffer ML, Garman RD, Greenstein JL, Juo M, Rogers BL;

XX Brauer AW;

XX WPI; 1991-164136/22.

XX N-PSDB; AAQ11838.

XX New pure covalently linked human T cell reactive feline protein -
PT and modified peptide(s), used to reduce effects of cat allergens
PT and to diagnose sensitivity to allergens.

XX Claim 2; Fig 3; 70pp; English.

XX Poly-A mRNA from cat parotid and mandibular glands was used to
CC produce cDNA clones for both chain 1 and chain 2 of TRFP. These
CC clones were then used to screen a cat genomic library. Chain 1
CC exists in two forms having different leader sequences (A and B).
CC The sequence can be used to express the protein and peptide derivs.
CC which stimulate T-cells in persons allergic to cats. The peptides
CC can be used to reduce/eliminate the allergic response partic. by

CC modifcn. of lymphokine prodn. by the T-cells. They can also be
 CC used to identify epitopes responsible for sensitivity. The DNA can
 CC be used to detect comparable sequence in other species, and also
 CC for prodn. of modified forms of TRFP esp. showing reduced binding
 CC to IgE and thus reduced tendency to cause adverse reactions.
 CC See also AAR12119-R12123.

CC Sequence 111 AA:

Query Match 89.2%; Score 99; DB 12; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.2e-88;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DTRMGALVIALLVLTQALGVMAETCPFIYDFVFAVANGNELLDLSLTFKVATEPERTA 60
 DB 1 dtrmgallviallvltqalgvmaetcpfiydfvfvangnellldlsltkvatpertata 60
 OY 61 MKKIDQVYENGILSRVLDGLVMTTSSSKDCMGAEVON 99
 DB 61 mkkidqvyengilstrvldglvmttssskdcmgaeavn 99

RESULT 10
 AAB28937 standard; Protein: 92 AA.

AC AAB28937;

DT 29-JAN-2001 (first entry)

DE T cell reactive feline protein chain 2 PRO long.

KW Cat; allergy; human T cell reactive feline protein; hTRFP;

KM Immunotherapy.

OS Fells sp.

PN US6120769-A.

PD 19-SEP-2000.

PF 28-APR-1995; 95US-0431184.

PR 02-SEP-1994; 94US-0300928.

PR 03-NOV-1989; 89US-0431565.

PR 28-FEB-1991; 91US-0662276.

PR 13-DEC-1991; 91US-0807529.

PR 25-MAR-1992; 92US-0857311.

PR 15-MAY-1992; 92US-0884718.

PR 15-JAN-1993; 93US-0006116.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Gelfer ML, Garman RD, Greenstein JL, Bond JF;

PT WPI: 2000-601477/57.

DR Detecting, preventing and treating sensitivity to cat protein allergen

PT comprises combining a biological sample with a human T cell reactive

PT feline protein and determining the extent of binding that occurs

XX Disclosure: Figure 7; 106pp; English.

XX The present invention relates to the detection of sensitivity to a cat

XX protein allergen by combining a blood sample from a subject with a

XX peptide of human T cell reactive feline protein (hTRFP). This method

XX and the hTRFP peptides are useful for diagnosing, preventing and

XX treating cat allergies by reducing or abolishing an individual's

XX allergic response to a cat allergen. DNA encoding the TRFP may be

CC functional equivalents useful in immunotherapy. The present
 CC sequence was used in the invention;
 CC
 XX Sequence 92 AA:

Query Match 82.9%; Score 92; DB 21; Length 92;
 Best Local Similarity 100.0%; Pred. No. 6.1e-82;
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VMMAETCPFIYDFVFAVANGNELLDLSLTFKVATEPERTAMKKTQDQYENGILSRVLD 79
 DB 1 vmmaetcpfiydfvfvangnellldlsltkvatpertamkkidqvyengilstrvld 60
 OY 80 GLVMTTSSSKDCMGAEVONTEDIKMTIGR 111
 DB 61 glvmttssskdcmgaeavntedikmtlgr 92

RESULT 11
 AAY87678 standard; Protein: 92 AA.

AC AAY87678;

DT 22-AUG-2000 (first entry)

DE Feline human TRFP chain 2 long form protein #2.

KW T-cell reactive feline protein; TRFP; Fel d I; cat allergen;

KM antiallergic; T cell stimulator; diagnostic; immunotherapy.

OS Fells sp.

PN US6048962-A.

PD 11-APR-2000.

PF 27-APR-1995; 95US-0430014.

PR 02-SEP-1994; 94US-0300928.

PR 03-NOV-1989; 89US-0431565.

PR 28-FEB-1991; 91US-0662276.

PR 13-DEC-1991; 91US-0807529.

PR 25-MAR-1992; 92US-0857311.

PR 15-MAY-1992; 92US-0884718.

PR 15-JAN-1993; 93US-0006116.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Kuo M, Rogers BL, Gelfer ML, Morgenstern JP, Brauer AW,

PT Greenstein JL, Griffith IJ, Garman RD;

PT WPI: 2000-316905/27.

DR New human T cell reactive feline protein useful for reducing or

PT abolishing individual's allergic response to cat allergen comprising

PT two different covalently linked peptide chains

XX Example 2; Column 81-84; 106pp; English.

XX This invention describes a novel naturally occurring cat protein allergen

XX (I), human T cell reactive feline protein (TRFP), comprising two

XX different covalently linked peptide chains with a molecular weight of 20

XX kD, 40 kD or 130 kD under non-reducing conditions and 5 kD or 10-18 kD

XX under reducing conditions. The products of the invention have

XX antiallergic activity and act as human T cell stimulators. TRFP is useful

XX for reducing or preventing the adverse effects of cat allergens on cat

CC immunotherapy against cat allergy. DNA sequences encoding TRFP are
 CC useful as probes to locate equivalent sequences present in other species
 CC (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics
 CC and/or therapeutics. Fully defined and characterized TRFP provides
 CC complete and a very simple desensitization therapy. This sequence
 CC represents a human T cell reactive feline protein (also known as Fel d 1)
 CC chain 2, long form which is described in the method of the invention.

XX Sequence 92 AA:

Query Match 82.9%; Score 92; DB 21; Length 92;
 Best Local Similarity 100.0%; Pred. No. 6.1e-82;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMAETCPFFYDVFVAVANGNELLDLSITKYNATEPERTAMKIKDDCYVENGILSRVID 79
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 vkmaetcpffydvvfvaangnellldlsitkynatepertamkikddcyvengilsrvid 60

QY 80 GLVMTTSSSKDGMGEAVONTVEDLKLNTLGR 111
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 glvmttssskdgmgeavontvedlklntlgr 92

RESULT 12

AAV51475
 ID AAV51475 standard; Protein; 92 AA.

AC AAV51475;

DT 22-MAY-2000 (first entry)

DE Human TRFP chain 2 (long form) protein fragment #2.

KW T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;
 KW down regulation; immune response; allergen; immunoglobulin E;
 OS sensitivity; cat protein allergen; human; chain 2.

XX Homo sapiens.

PN US6019972-A.

PD 01-FEB-2000.

PF 02-SEP-1994; 94US-0300928.

PR 03-NOV-1989; 89US-0431565.

PR 28-FEB-1991; 91US-0662276.

PR 13-DEC-1991; 92US-0807529.

PR 25-MAR-1992; 92US-0857311.

PR 15-MAY-1992; 92US-0884718.

PR 15-JAN-1993; 93US-0006116.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gelfer ML;

PT WPI: 2000-146862/13.

XX peptides of human T cell reactive feline protein for treating

XX sensitivity to cat protein allergens comprise at least one T cell

XX epitope recognized by a T cell receptor specific for the human T cell

XX reactive feline protein -

XX Example 1; Column 83-84; 105pp; English.

XX This invention describes a novel peptide (I) of human T cell reactive

XX feline protein (TRFP) having at least one T cell epitope recognized

XX by a T cell receptor specific for the human T cell reactive feline

XX protein, the peptide consisting of at least 7-30 amino acids, and having

XX an amino acid sequence derived from an amino acid sequence comprising 94,

XX 96, 97, 109, or 111 residues, given in the specification. The peptides

XX down regulate the immune response to the allergen. The peptides have

CC reduced immunoglobulin E binding and reduce T cell responsiveness. The
 CC peptide (I) is useful in compositions for treating sensitivity to a cat
 CC protein allergen in a subject. This sequence represents the human TRFP
 CC chain 2 (long form).

XX Sequence 92 AA:

Query Match 82.9%; Score 92; DB 21; Length 92;
 Best Local Similarity 100.0%; Pred. No. 6.1e-82;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMAETCPFFYDVFVAVANGNELLDLSITKYNATEPERTAMKIKDDCYVENGILSRVID 79
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 vkmaetcpffydvvfvaangnellldlsitkynatepertamkikddcyvengilsrvid 60

QY 80 GLVMTTSSSKDGMGEAVONTVEDLKLNTLGR 111
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 61 glvmttssskdgmgeavontvedlklntlgr 92

RESULT 13

AAK12123
 ID AAK12123 standard; Protein; 97 AA.

AC AAK12123;

DT 26-JUL-1991 (first entry)

DE TRFP chain 2 - truncated short form.

KW Human T cell reactive feline protein; cat allergens.

XX Felis catus.

OS Key Location/Qualifiers

FT Peptide 3..19

FT Protein /label= Leader sequence

FT /label= TRFP I truncated short form

PN WO9106571-A.

PD 16-MAY-1991.

PF 02-NOV-1990; 90MO-US06548.

PR 03-NOV-1989; 89US-0431565.

PA (IMMU-) IMMUNOLOGIC PHARM COR.

PI Gelfer ML, Garman RD, Greenstein JL, Juo M, Rogers BL;

PT Brauer AW;

DR WPI: 1991-164136/22.

XX N-PSDB; AAK11840.

XX New pure covalently linked human T cell reactive feline protein -

XX and modified peptide(s), used to reduce effects of cat allergens

XX and to diagnose sensitivity to allergens.

XX Claim 2; Fig 5; 70pp; English.

XX Poly-A mRNA from cat parotid and mandibular glands was used to

XX produce cDNA clones for both chain 1 and chain 2 of TRFP. These

XX clones were then used to screen a cat genomic library. Chain 1

XX exists in two forms having different leader sequences (A and B).

XX The sequence can be used to express the protein and peptide derivs.

XX which stimulate T-cells in persons allergic to cats. The peptides

XX can be used to reduce/eliminate the allergic response partic. by

XX modifcn. of lymphokine prodn. by the T-cells. They can also be

XX used to identify epitopes responsible for sensitivity. The DNA can

XX be used to detect comparable sequence in other species, and also

CC useful as probes to locate equivalent sequences present in other species
CC (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics
CC and/or therapeutics. Fully defined and characterized TRP provides
CC complete and a very simple desensitization therapy. This sequence
CC represents a human T cell reactive feline protein (also known as Fel d 1)
CC chain 2, short form which is described in the method of the invention.
XX
SQ Sequence 109 AA:

	74.8%;	Score 83;	DB 21;	Length 109;
Query Match	Similarity	100.0%;	Pred. NO.	4.1e-73;
Best Local	Conservative	0;	Mismatches	0;
Matches	Indels	Gaps		
OY	1 DDMRALLVIALVYQALGVKMAECPPIFYDVFPFVANGNELLDLSLKVNAPEPERA	60		
db	1 ddmrallviallvqalgvkmaecpfiydfavangnellldslkvnapertpa	60		
OY	61 MKRIODCYENGILSRVLDTLVN	83		
db	61 mkriodcyengilsvrlvdglvm	83		

Search completed: May 7, 2002, 12:17:34
Job time: 201 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:10:08 ; Search time 23.73 seconds
(without alignments)
346.487 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 532
Sequence: 1 DTMRGALLVALLVLTQALGV.....CMGEAVNTVEDLKLNTLGR 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.1101.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	552	100.0	111	13	TRFP Chain #2 with
2	552	100.0	111	21	T cell reactive fe
3	552	100.0	111	21	Feline human TRFP
4	552	100.0	111	21	Cat TRFP chain 2 1
5	552	100.0	111	21	Human TRFP chain 2
6	547	99.1	111	12	TRFP chain 2 - lon
7	541	98.0	109	14	Human T cell react
8	541	98.0	109	14	Felis sp. allergen
9	541	98.0	109	20	TRFP Chain #2 with
10	512.5	92.8	110	13	T cell reactive fe
11	506	91.7	109	21	AA828934

12	506	91.7	109	21	AA87674
13	506	91.7	109	21	AA890104
14	506	91.7	109	21	AA8901471
15	498	90.2	109	12	AA812122
16	465	84.2	92	21	AA828937
17	465	84.2	92	21	AA87678
18	465	84.2	92	21	AA851475
19	458	83.0	92	18	AA827382
20	458	83.0	92	20	AA825532
21	457.5	82.9	101	13	AA827372
22	419	75.9	90	21	AA828938
23	419	75.9	90	21	AA87679
24	419	75.9	90	21	AA851476
25	412	74.6	97	12	AA812123
26	408	73.9	97	21	AA828935
27	408	73.9	97	21	AA87675
28	408	73.9	97	21	AA890105
29	408	73.9	97	21	AA851472
30	370.5	67.1	82	21	AA828940
31	370.5	67.1	82	21	AA87681
32	370.5	67.1	82	21	AA890106
33	370.5	67.1	82	21	AA851478
34	288.5	52.3	82	13	AA827371
35	185	33.5	51	21	AA828939
36	185	33.5	51	21	AA87680
37	185	33.5	51	21	AA851477
38	128.5	23.3	96	14	AA836548
39	128.5	23.3	96	21	AA828979
40	128.5	23.3	96	21	AA890145
41	128	23.2	26	14	AA836544
42	128	23.2	26	14	AA841977
43	128	23.2	26	21	AA828943
44	128	23.2	26	21	AA87684
45	128	23.2	26	21	AA890109

ALIGNMENTS

RESULT 1
ID AAR27369 standard; protein; 111 AA.
AC AAR27369;
XX
DT 20-MAY-1998 (first entry)
XX
DE TRFP Chain #2 with C2 leader sequence.
XX
KW T cell reactive feline protein; cat allergy; allergic; Ige;
desensitizing.
XX
OS Felis domesticus.
XX
FH Key
FT Peptide
FT protein
FT
PN WO9215613-A.
PD 17-SEP-1992.
PF 20-FEB-1992; 92WO-US01344.
PR 28-FEB-1991; 91US-0662193.
PA (IMMUNO-) IMMUNOLOGIC PHARM CORP.
PI Bond J, Kuo M;
XX
DR WPI; 1992-331670/40.

Feline human TRFP
Cat TRFP chain 2 s
Human TRFP chain 2
TRFP I chain 2 - s
T cell reactive fe
Feline human TRFP
Human TRFP chain 2
Human 2 of major c
Human MHC class II
TRFP chain #2 C2ST
T cell reactive fe
Feline human TRFP
Human TRFP chain 2
TRFP chain 2 - tru
T cell reactive fe
Feline human TRFP
Cat TRFP chain 2 t
Human TRFP chain 2
T cell reactive fe
Feline human TRFP
Human TRFP chain 2
TRFP chain 2 - tru
T cell reactive fe
Feline human TRFP
Cat TRFP chain 2
Human TRFP chain 2
TRFP chain #2 C2ST
T cell reactive fe
Feline human TRFP
Human TRFP chain 2
Recombinant YZX.
Peptide YZX. Feli
TRFP fusion protei
Peptide 2. Felis.
Human T cell react
Peptide 2 derived
Feline human TRFP
Cat TRFP derived p

XX Modified human T-cell reactive feline protein - stimulates T-cell
PT In individuals allergic to cats and shows reduced
PR histamine-releasing properties
PS
XX Claim 1: Fig 1: 35pp; English.
XX
CC This sequence represents a modified human T-cell reactive feline
CC protein which stimulates T-cells from an individual who is allergic
CC to cats, but which interacts with human IgE to a lesser extent than
CC does affinity purified TRFP. The protein is modified by treating
CC with either a mild alkali (pH 12.5-13.5, KOH, NaOH, LiOH or tertiary
CC amines) or an enzyme which removes O-linked groups (carbohydrate
CC moieties). It is useful in desensitising people who are allergic to cats.
XX
SQ Sequence 111 AA:

Query Match 100.0%; Score 552; DB 13; Length 111;
Best Local Similarity 100.0%; Pred. No. 2e-57;
Matches 111: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTRGALLVALLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLRVNAPEPRTA 60
Db 1 dtrgallvlllvtqalgvkmaetcpfydvfpavangnellldlslrkvatepeta 60
Qy 61 MKRIQCYVENCGLSRVLDGLVMTTSSSKCKMGEAVQNTVEDLKLNTLGR 111
Db 61 mkrldcyvenglsrvldglvmttssskckmgeavqntvedlklntlgr 111

RESULT 2

AA828933 ID AAB28933 standard; Protein: 111 AA.

XX AAB28933:

XX 29-JAN-2001 (first entry)

XX T cell reactive feline protein chain 2 long form.

XX Cat: allergy; human T cell reactive feline protein; hTRFP;

XX Immunotherapy.

XX Fells sp.

XX US6120769-A.

XX 19-SEP-2000.

XX 28-APR-1995; 95US-0431184.

XX 02-SEP-1994; 94US-0300928.

XX 03-NOV-1989; 89US-0431565.

XX 28-FEB-1991; 91US-0662276.

XX 13-DEC-1991; 91US-0807529.

XX 25-MAR-1992; 92US-0857311.

XX 15-MAY-1992; 92US-0884718.

XX 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Geffer ML, Garman RD, Greenstein JL, Bond JF;

XX WPI: 2000-601477/57.

XX N-PSDB: AAC60103.

XX Detecting, preventing and treating sensitivity to cat protein allergen

XX comprises combining a biological sample with a human T cell reactive

XX feline protein and determining the extent of binding that occurs -

XX Claim 1: Figure 3: 106pp; English.

CC The present invention relates to the detection of sensitivity to a cat
CC protein allergen by combining a blood sample from a subject with a
CC peptide of human T cell reactive feline protein (hTRFP). This method
CC and the hTRFP peptides are useful for diagnosing, preventing and
CC treating cat allergies by reducing or abolishing an individual's
CC allergic response to a cat allergen. DNA encoding the TRFP may be
CC used as probes to locate equivalent sequences present in other species.
CC These may further be used to study the mechanism of immunotherapy of
CC cat allergy, and to design modified derivatives, analogues or
CC functional equivalents useful in immunotherapy. The present
CC sequence was used in the invention.
XX
SQ Sequence 111 AA:

Query Match 100.0%; Score 552; DB 21; Length 111;
Best Local Similarity 100.0%; Pred. No. 2e-57;
Matches 111: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DTRGALLVALLVLTQALGVKMAETCPFYDVFPAVANGNELLDLSLRVNAPEPRTA 60
Db 1 dtrgallvlllvtqalgvkmaetcpfydvfpavangnellldlslrkvatepeta 60
Qy 61 MKRIQCYVENCGLSRVLDGLVMTTSSSKCKMGEAVQNTVEDLKLNTLGR 111
Db 61 mkrldcyvenglsrvldglvmttssskckmgeavqntvedlklntlgr 111

RESULT 3

AA87673 ID AAY87673 standard; Protein: 111 AA.

XX AAY87673:

XX 22-AUG-2000 (first entry)

XX Feline human TRFP chain 2 long form protein.

XX T-cell reactive feline protein; TRFP; Fel d I; cat allergen;

XX anti-allergic; T cell stimulator; diagnostic; immunotherapy.

XX Fells sp.

XX US6048962-A.

XX 11-APR-2000.

XX 27-APR-1995; 95US-0430014.

XX 02-SEP-1994; 94US-0300928.

XX 03-NOV-1989; 89US-0431565.

XX 28-FEB-1991; 91US-0662276.

XX 13-DEC-1991; 91US-0807529.

XX 25-MAR-1992; 92US-0857311.

XX 15-MAY-1992; 92US-0884718.

XX 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMULOGIC PHARM CORP.

XX Kuo M, Rogers BL, Geffer ML, Morgenstern JP, Brauer AW;

XX Greenstein JL, Griffith ID, Garman RD;

XX WPI: 2000-316905/27.

XX N-PSDB: AAA12244.

XX New human T cell reactive feline protein useful for reducing or

XX abolishing individual's allergic response to cat allergen comprising

XX two different covalently linked peptide chains -

XX Claim 2: Column 75-76; 106pp; English.

CC This invention describes a novel naturally occurring cat protein allergen
CC (I), human T cell reactive feline protein (TRFP), comprising two

CC different covalently linked peptide chains with a molecular weight of 20
 CC kD, 40 kD or 130 kD under non-reducing conditions and 5 kD or 10-18 kD
 CC under reducing conditions. The products of the invention have
 CC antiallergic activity and act as human T cell stimulators. TRFP is useful
 CC for reducing or preventing the adverse effects of cat allergens on cat
 CC allergic individuals and in ex vivo diagnostic tests to determine which
 CC peptides cause sensitivity so as to selectively use them to desensitize
 CC a cat sensitive individual. Purified TRFP is also useful for studying
 CC the mechanism of immunotherapy of cat allergy and to design modified
 CC derivatives, analogs or functional equivalents that are more useful in
 CC immunotherapy against cat allergy. DNA sequences encoding TRFP are
 CC useful as probes to locate equivalent sequences present in other species
 CC (goats, sheep, dogs, rabbits or horses) that may be useful in diagnostics
 CC and/or therapeutics. Fully defined and characterized TRFP provides
 CC complete and a very simple desensitization therapy. This sequence
 CC represents a human T cell reactive feline protein (also known as Fel d 1)
 CC chain 2, long form which is described in the method of the invention.
 CC
 XX
 SQ Sequence 111 AA:

Query Match 100.0%; Score 552; DB 21; Length 111;
 Best Local Similarity 100.0%; Pred. No. 2e-57; Mismatches 0; Indels 0; Gaps 0;
 Matches 111; Conservative 0;

QY 1 DPMRGALVLAIVLVTQALGVKMAETCPIFYDVFAVANGNELLDLSLTKVNATEPERTA 60
 Db 1 dcmrgallvllalvtqalgvkmactcpifydvffavangnellldsltkvnatepercta 60
 QY 61 MKKIDCVYENGILSRVLDGLVFTTSSSKDCGAEAVQNTVEDLKINTLGR 111
 Db 61 mkkidgyengilsrvldglvmtltssskdcgaeavqntvedlkinltgr 111

RESULT 4
 AAY90103 standard; Protein: 111 AA.
 XX AAY90103;

-DT 13-JUL-2000 (first entry)

DE Cat TRFP chain 2 long form protein sequence.

KW Cat: TRFP: human T-cell reactive feline protein; cat protein allergen;
 -KW house dust; Fel d 1; cat allergy; Felis domesticus sensitivity; therapy;
 KW diagnosis; goat; sheep; horse; rabbit; dog.

OS Felis domesticus.

XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT Protein /note= "signal peptide"
 FT /note= "mature TRFP chain 2 long form"

PN US6025162-A.

PD 15-FEB-2000.

PF 28-APR-1995; 95US-0430944.

PR 02-SEP-1994; 94US-0300928.
 PR 03-NOV-1989; 89US-0431565.
 PR 28-FEB-1991; 91US-0662276.
 PR 13-DEC-1991; 91US-0807529.
 PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 PA
 PI Morgenstern JP, Griffith ID, Rogers BL;

XX
 DR WPI: 2000-181812/16.
 DR N-PSDB; AAA07437.
 PT New human T cell reactive feline protein, useful for desensitizing cat
 PT allergic individuals to cat allergens -
 PS
 XX Claim 1: Fig 3; 108pp; English.

CC This sequence is a peptide chain of the human T cell reactive feline
 CC protein (TRFP) of the invention. The protein is a cat protein allergen,
 CC and was isolated from a vacuum bag extract obtained by affinity
 CC purification of house dust collected from several homes with cats. TRFP
 CC is composed of two covalently linked peptide chains, and is also referred
 CC to as Fel d 1. TRFP and its peptides are useful for reducing or
 CC preventing the adverse effects that exposure to cat allergens normally
 CC has on cat allergic individuals (i.e. to desensitize individuals to cat
 CC allergens or block the effect of the allergens). TRFP is also used in
 CC methods of diagnosing sensitivity to Felis domesticus in an individual.
 CC DNA sequences encoding TRFP can be used as probes to locate equivalent
 CC sequences present in other species, e.g. goat, sheep, horse, rabbit and
 CC dog, that may be useful in a diagnostic and/or therapeutic applications.
 CC
 XX
 SQ Sequence 111 AA:

Query Match 100.0%; Score 552; DB 21; Length 111;
 Best Local Similarity 100.0%; Pred. No. 2e-57; Mismatches 0; Indels 0; Gaps 0;
 Matches 111; Conservative 0;

QY 1 DPMRGALVLAIVLVTQALGVKMAETCPIFYDVFAVANGNELLDLSLTKVNATEPERTA 60
 Db 1 dcmrgallvllalvtqalgvkmactcpifydvffavangnellldsltkvnatepercta 60
 QY 61 MKKIDCVYENGILSRVLDGLVFTTSSSKDCGAEAVQNTVEDLKINTLGR 111
 Db 61 mkkidgyengilsrvldglvmtltssskdcgaeavqntvedlkinltgr 111

RESULT 5
 AAY51470 standard; Protein: 111 AA.
 XX AAY51470;

-DT 22-MAY-2000 (first entry)

DE Human TRFP chain 2 (long form) protein fragment.

KW T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;
 -KW down regulation; immune response; allergen; immunoglobulin E;
 KW sensitivity; cat protein allergen; human; chain 2.

OS Homo sapiens.

PN US6019972-A.

PD 01-FEB-2000.

PF 02-SEP-1994; 94US-0300928.

PR 03-NOV-1989; 89US-0431565.
 PR 28-FEB-1991; 91US-0662276.
 PR 13-DEC-1991; 91US-0807529.
 PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.

XX (IMMU-) IMMUNOLOGIC PHARM CORP.
 PA
 PI Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gelfer ML;
 XX WPI: 2000-146862/13.

DR N-PSDB; AA288617.
 XX Peptides of human T cell reactive feline protein for treating
 PT sensitivity to cat protein allergens comprise at least one T cell
 PT epitope recognized by a T cell receptor specific for the human T cell
 PS reactive feline protein -
 XX
 PS Claim 1: Column 75-78; 105pp; English.
 CC This invention describes a novel peptide (I) of human T cell reactive
 CC feline protein (hTRFP) having at least one T cell epitope recognized
 CC by a T cell receptor specific for the human T cell reactive feline
 CC protein, the peptide consisting of at least 7-30 amino acids, and having
 CC an amino acid sequence derived from an amino acid sequence comprising 94,
 CC 96, 97, 109, or 111 residues, given in the specification. The peptides
 CC down regulate the immune response to the allergen. The peptides have
 CC reduced immunoglobulin E binding and reduce T cell responsiveness. The
 CC peptide (I) is useful in compositions for treating sensitivity to a cat
 CC protein allergen in a subject. This sequence represents the human TRFP
 CC chain 2 (long form).
 XX
 SQ Sequence 111 AA:
 Query Match 100.0%; Score 552; DB 21; Length 111;
 Best Local Similarity 100.0%; Pred. No. 26-57;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 -QY 1 DTRMGALVLTALVTQALGVKMAETCPFEYDVFPAVANGNELLDLSTKYNATEPERTA 60
 DB 1 dtrmgallvltallvtqalgvykmaetcpfydvffavangnellldstlkyateperpta 60
 QY 61 MKRIQCYVENGILSRVLQGLVMTTSSSKDCMGCAVQNTVEDLKINTLGR 111
 DB 61 mkriqcyvengilsvrlqglvmttssskdcmgcaqvntvedlkinltlgr 111
 Db 61 mkriqcyvengilsvrlqglvmttssskdcmgcaqvntvedlkinltlgr 111
 RESULT 6
 AAR12121
 ID AAR12121 standard; Protein: 111 AA.
 AC AAR12121;
 XX
 DT 26-JUL-1991 (first entry)
 XX
 DE TRFP chain 2 - long form.
 XX
 KW Human T cell reactive feline protein; cat allergens.
 XX
 OS Felis catus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 3..19
 FT /label- leader sequence
 FT 20..111
 FT /label- TRFP I long form
 XX
 PN WO9106571-A.
 XX
 PD 16-MAY-1991.
 XX
 PF 02-NOV-1990; 90WO-US06548.
 XX
 PR 03-NOV-1989; 89US-0431565.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM COR.
 XX
 PI Geffer ML, Garman RD, Greenstein JL, Uno M, Rogers BL;
 PI Brauer AW;
 XX
 DR WPI: 1991-164136/22.
 DR N-PSDB; AAQ11838.
 XX

PT New pure covalently linked human T cell reactive feline protein -
 PT and modified peptide(s), used to reduce effects of cat allergens
 PT and to diagnose sensitivity to allergens.
 XX
 PS Claim 2: Fig 3; 70pp; English.
 CC Poly-A mRNA from cat parotid and mandibular glands was used to
 CC produce cDNA clones for both chain 1 and chain 2 of TRFP. These
 CC clones were then used to screen a cat genomic library. Chain 1
 CC exists in two forms having different leader sequences (A and B).
 CC The sequence can be used to express the protein and peptide deriva-
 CC which stimulate T-cells in persons allergic to cats. The peptides
 CC can be used to reduce/eliminate the allergic response partic. by
 CC modification of lymphokine prod. by the T-cells. They can also be
 CC used to identify epitopes responsible for sensitivity. The DNA can
 CC be used to detect comparable sequence in other species, and also
 CC for prod. of modified forms of TRFP esp. showing reduced binding
 CC to IgE and thus reduced tendency to cause adverse reactions.
 CC See also AAR12119-R12123.
 XX
 SQ Sequence 111 AA:
 Query Match 99.1%; Score 547; DB 12; Length 111;
 Best Local Similarity 99.1%; Pred. No. 7,86-57;
 Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DTRMGALVLTALVTQALGVKMAETCPFEYDVFPAVANGNELLDLSTKYNATEPERTA 60
 DB 1 dtrmgallvltallvtqalgvykmaetcpfydvffavangnellldstlkyateperpta 60
 QY 61 MKRIQCYVENGILSRVLQGLVMTTSSSKDCMGCAVQNTVEDLKINTLGR 111
 DB 61 mkriqcyvengilsvrlqglvmttssskdcmgcaqvntvedlkinltlgr 111
 Db 61 mkriqcyvengilsvrlqglvmttssskdcmgcaqvntvedlkinltlgr 111
 RESULT 7
 AAR36541
 ID AAR36541 standard; Protein: 109 AA.
 AC AAR36541;
 XX
 DT 12-AUG-1993 (first entry)
 XX
 DE TRFP chain 2 (with leader).
 XX
 KW Human T cell reactive feline protein; TRFP; leader A; leader B;
 XX
 OS Felis.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT /label- leader-peptide
 XX
 PN WO9308280-A.
 XX
 PD 29-APR-1993.
 XX
 PF 16-OCT-1992; 92WO-US08694.
 XX
 PR 16-OCT-1991; 91US-0777859.
 PR 13-DEC-1991; 91US-0807529.
 XX
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Kuo M, Morgenstern JP, Morville M;
 PI Rogers BL;
 XX
 DR WPI: 1993-152473/18.
 DR N-PSDB; AAQ41558.
 XX
 PT Recombinote peptide having T-cell stimulating activity - for the

PT diagnosis and treatment of sensitivity to protein allergens,
 PT auto:antigens and protein antigens
 XX
 PS Disclosure: Fig 2; 73pp; English.
 XX
 CC Chains 1 and 2 of the TRFP have been recombinantly expressed in E.
 CC coli and purified. T cell epitope studies using overlapping peptide
 CC regions derived from the TRFP amino acids sequence were used to
 CC identify multiple T cell epitopes in each chain of TRFP.
 XX
 SQ Sequence 109 AA:

Query Match 98.0%; Score 541; DB 14; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.9e-56;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MEGALLVLLVLTQALGVMAETCPFYDFVFAVANGNELLDLSITKYNATEPERTAK 62
 DB 1 mrgallvllalvlgqalgvmaetcpfydvffavangnellldlsitkynatepertank 60
 OY 63 KIODCYVENGILSRVLDGLVMTTSSSKDCMGCAVONTVEDLKINTLGR 111
 DB 61 kiodcyvengilstrvldglvmttssskdcmgcaavontvedlkintlgr 109

RESULT 8

AAR41985
 ID AAR41985 standard; protein: 109 AA.

AC AAR41985;
 XX

DT 21-APR-1994 (first entry)
 XX

DE Human T cell reactive feline protein chain 2.
 XX

KW Human; T cell; reactive; feline; protein; immune response; antigen;
 KW Crytomeria; Alternaria; Alder; Betula; Quercus; Olea; Artemesia;
 KW Plantago; Parietaria; Blatella; Apis; Periplaneta; autoantigen.
 XX

OS Homo sapiens.
 XX

Key Location/Qualifiers
 FH 1.17
 FT /note="Signal peptide"
 FT 18.109
 FT Protein /note="Mature protein"

WO9319178-A.
 XX

PD 30-SEP-1993.
 XX

PF 25-MAR-1993; 93WO-US02462.
 XX

PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.
 XX

PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX

PI Briner TJ, Garman RD, Geffer ML, Greenstein JL, Kuo M;
 PI Morville M;
 XX

DR WPI; 1993-320744/40.
 DR N-PSDB; AAQ49535.
 XX

PT New peptide(s) for inducing tolerance - comprise one or more
 PT epitope(s) of an allergen administered subcutaneously, for
 PT treating sensitivity to cats, bees, etc.
 XX

PS Disclosure: Fig 2; 107pp; English.
 XX

CC This sequence represents chain 2 of human T cell reactive feline
 CC protein (TRFP). Peptides derived from TRFP may be used in a
 CC therapeutic composition which is useful in treating diseases which
 CC involve an immune response to a protein antigen. This composition
 CC may be used to induce tolerance in a mammal to Dermatophagoides,
 CC Fells, Ambrosia, Lolium, Cryptomeria, Alternaria, Alder, Betula,
 CC Quercus, Olea, Artemesia, Plantago, Parietaria, Canis, Blatella,
 CC Apis, Periplaneta and to autoantigens in humans.
 XX
 SQ Sequence 109 AA:

Query Match 98.0%; Score 541; DB 14; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.9e-56;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MEGALLVLLVLTQALGVMAETCPFYDFVFAVANGNELLDLSITKYNATEPERTAK 62
 DB 1 mrgallvllalvlgqalgvmaetcpfydvffavangnellldlsitkynatepertank 60
 OY 63 KIODCYVENGILSRVLDGLVMTTSSSKDCMGCAVONTVEDLKINTLGR 111
 DB 61 kiodcyvengilstrvldglvmttssskdcmgcaavontvedlkintlgr 109

RESULT 9

AAZ5593
 ID AAZ5593 standard; protein: 109 AA.

AC AAZ5593;
 XX

DT 30-SEP-1999 (first entry)
 XX

DE Fells sp. allergen 1082946 Fel di chain 2 precursor protein fragment.
 XX

KW Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; mould; food; insect; sting;
 KW chironomidae; spider; mite; housefly; fruit fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mite; gerbil; vaccine; treatment; prevention; hypersensitivity.
 XX

OS Fells sp.
 XX

PN WO9934826-A1.
 XX

PD 15-JUL-1999.
 XX

PF 11-JAN-1999; 99WO-GB00080.
 XX

PR 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX

PI Kay AB, Larche M;
 XX

DR WPI; 1999-458255/38.
 XX

PT Desensitizing patients to polypeptide allergens
 XX

Example 6; Page 53; 117pp; English.
 XX

CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,

CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrink mottle beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents the cat allergen 1082946 Fel d1 chain 2 precursor.
 XX
 SQ Sequence 109 AA;

Query Match 98.0%; Score 541; DB 20; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.9e-56;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVALLVLTQALGVKMAETCPFYDVEFAVANGNELLDLSITRVNATEPPTAK 62
 DB 1 mrgallviallvltqalgvkmaetcpfydvifavangnelldlsitrvnateppta 60
 QY 63 KIQDCYVENGLISRVLDGLVMTTSSSKDCGAEAVONTVEDIKLNTLGR 111
 DB 61 kiqdcyvenglisrvldglvmttssskdcgaeavontvediklntlgr 109

RESULT 10
 AAR27370
 ID AAR27370 standard; protein; 110 AA.
 AC AAR27370;
 DT 20-MAY-1998 (first entry)
 DE TRFP Chain #2 with C2 leader sequence.
 KW T cell reactive feline protein; cat allergy; allergic; IGE;
 KM desensitizing.
 OS Fells domesticus.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT Label- C2 leader
 FT protein 19..110
 FT Label- TRFP chain #2
 PA MO9215613-A.
 PD 17-SEP-1992.
 PF 20-FEB-1992; 92MO-US01344.
 PR 28-FEB-1991; 91US-0662193.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond J, Kuo M;
 PS WPI; 1992-331670/40.
 PT Modified human T-cell reactive feline protein - stimulates T-cell
 PT in individuals allergic to cats and shows reduced
 PT histamine-releasing properties
 PS Claim 1; Fig 1; 35pp; English.
 CC This sequence represents a modified human T-cell reactive feline
 CC protein which stimulates T-cells from an individual who is allergic
 CC to cats, but which interacts with human IGE to a lesser extent than
 CC does affinity purified TRFP. The protein is modified by treating
 CC with either a mild alkali (pH 12.5-13.5, KOH, NaOH, LiOH or tertiary
 CC amines) or an enzyme which removes O-linked groups (carbohydrate
 CC moieties). It is useful in desensitizing people who are allergic to cats.
 XX
 SQ Sequence 110 AA;

Query Match 92.8%; Score 512.5; DB 13; Length 110;
 Best Local Similarity 94.6%; Pred. No. 8.9e-53;
 Matches 105; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 DTRGALLVALLVLTQALGVKMAETCPFYDVEFAVANGNELLDLSITRVNATEPPTAK 60
 DB 1 dtrgallviallvltqalgvkmaetcpfydvifavangnelldlsitrvnateppta 60
 QY 61 MKRIQDCYVENGLISRVLDGLVMTTSSSKDCGAEAVONTVEDIKLNTLGR 111
 DB 61 mkrkqdcyvenglisrvldglvmttssskdcgaeavontvediklntlgr 110

RESULT 11
 AAB28934
 ID AAB28934 standard; protein; 109 AA.
 AC AAB28934;
 DT 29-JAN-2001 (first entry)
 DE T cell reactive feline protein chain 2 short form.
 KW Cat; allergy; human T cell reactive feline protein; hTRFP;
 KM Immunotherapy.
 OS Fells sp.
 PN US6120769-A.
 PD 19-SEP-2000.
 PF 28-APR-1995; 95US-0431184.
 PR 02-SEP-1994; 94US-0300928.
 PR 03-NOV-1989; 89US-0431565.
 PR 28-FEB-1991; 91US-0662276.
 PR 13-DEC-1991; 91US-0807529.
 PR 25-MAR-1992; 92US-0857311.
 PR 15-MAY-1992; 92US-0884718.
 PR 15-JAN-1993; 93US-0006116.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Gelfer ML, Garman RD, Greenstein JL, Bond JF;
 DR WPI; 2000-601477/57.
 DR N-PSDB; AAC60104.
 PT Detecting, preventing and treating sensitivity to cat protein allergen
 PT comprises combining a biological sample with a human T cell reactive
 PT feline protein and determining the extent of binding that occurs -
 PS Claim ; Figure 4; 106pp; English.
 CC The present invention relates to the detection of sensitivity to a cat
 CC protein allergen by combining a blood sample from a subject with a
 CC peptide of human T cell reactive feline protein (hTRFP). This method
 CC and the hTRFP peptides are useful for diagnosing, preventing and
 CC treating cat allergies by reducing or abolishing an individual's
 CC allergic response to a cat allergen. DNA encoding the TRFP may be
 CC used as probes to locate equivalent sequences present in other species.
 CC These may further be used to study the mechanism of immunotherapy of
 CC cat allergy, and to design modified derivatives, analogues or
 CC functional equivalents useful in immunotherapy. The present
 CC sequence was used in the invention.
 XX
 SQ Sequence 109 AA;

Query Match 91.7%; Score 506; DB 21; Length 109;

CC has on cat allergic individuals (i.e. to desensitize individuals to cat
 CC allergens or block the effect of the allergens). TRFP is also used in
 CC methods of diagnosing sensitivity to feline allergens in an individual.
 CC DNA sequences encoding TRFP can be used as probes to locate relevant
 CC sequences present in other species, e.g. goat, sheep, horse, rabbit, and
 CC dog, that may be useful in a diagnostic and/or therapeutic applications.
 XX

Sequence 109 AA;

Query Match 91.7%; Score 506; DB 21; Length 109;
 Best Local Similarity 93.7%; Pred. No. 5,1e-52;
 Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 DTMGALLVALLVTOALGVKMAETCPFDVFFAVANGNELLDLSLRVNAITEPERTA 60
 Db 1 dtmgallviallvqalgvkmaetcpfdvffavangnelldlslrkvnatepeta 60
 QY 61 MKRTDQCYVENGISRVLDGLVNTTSSKDCGGEAVONTYEDLKLTGCR 111
 Db 61 mkrtidcyvenglstrvldglvnttsskdcggevontyvedlklntlgr 109

RESULT 14

ID AAY51471 standard; Protein; 109 AA.

AC AAY51471;

DM 22-MAY-2000 (first entry)

XX Human TRFP chain 2 (short form) protein fragment.

XX T-cell reactive feline protein; TRFP; T cell epitope; T cell receptor;
 XX down regulation; immune response; allergen; immunoglobulin E;
 XX sensitivity; cat protein allergen; human; chain 2.

OS Homo sapiens.

PN US6019972-A.

PN 01-FEB-2000.

PE 02-SEP-1994; 94US-0300928.

PE 03-NOV-1989; 89US-0431565.

PE 28-FEB-1991; 91US-0662276.

PE 13-DEC-1991; 91US-0857329.

PE 25-MAY-1992; 92US-0857321.

PE 15-MAY-1992; 92US-0884718.

PE 15-JAN-1993; 93US-0006116.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Garman RD, Greenstein JL, Kuo M, Briner TJ, Morville M, Gelfer ML;

PI WPI: 2000-146862/13.

DR N-PSDB: AA288618.

XX Peptides of human T cell reactive feline protein for treating

XX sensitivity to cat protein allergens comprised at least one T cell

XX epitope recognized by a T cell receptor specific for the human T cell

XX reactive feline protein -

XX Claim 1: Column 79-80; 105pp; English.

This invention describes a novel peptide (I) of human T cell reactive
 CC feline protein (TRFP) having at least one T cell epitope recognized
 CC by a T cell receptor specific for the human T cell reactive feline
 CC protein, the peptide consisting of at least 7-30 amino acids, and having
 CC an amino acid sequence derived from an amino acid sequence comprising 94,
 CC 96, 97, 109, or 111 residues, given in the specification. The peptides
 CC down regulate the immune response to the allergen. The peptides have

CC reduced immunoglobulin E binding and reduce T cell responsiveness. The
 CC peptide (I) is useful in compositions for treating sensitivity to a cat
 CC protein allergen in a subject. This sequence represents the human TRFP
 CC chain 2 (short form).
 XX

Sequence 109 AA;

Query Match 91.7%; Score 506; DB 21; Length 109;
 Best Local Similarity 93.7%; Pred. No. 5,1e-52;
 Matches 104; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 DTMGALLVALLVTOALGVKMAETCPFDVFFAVANGNELLDLSLRVNAITEPERTA 60
 Db 1 dtmgallviallvqalgvkmaetcpfdvffavangnelldlslrkvnatepeta 60
 QY 61 MKRTDQCYVENGISRVLDGLVNTTSSKDCGGEAVONTYEDLKLTGCR 111
 Db 61 mkrtidcyvenglstrvldglvnttsskdcggevontyvedlklntlgr 109

RESULT 15

ID AAR12122 standard; Protein; 109 AA.

AC AAR12122;

DM 26-JUL-1991 (first entry)

XX TRFP I chain 2 - short form.

XX Human T cell reactive feline protein; cat allergens.

XX Felis catus.

EH Key Location/Qualifiers

FT Peptide 3..19

FT Protein /Label- Leader sequence

FT 20..111 /Label- TRFP I short form

PN W09106571-A.

PN 16-MAY-1991.

PE 02-NOV-1990; 90MO-US06548.

PE 03-NOV-1989; 89US-0431565.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Gelfer ML, Garman RD, Greenstein JL, Juo M, Rogers BL;

PI Brauer AW;

PI WPI: 1991-164136/22.

DR N-PSDB: AAO11839.

XX New pure covalently linked human T cell reactive feline protein -

XX and modified peptide(s), used to reduce effects of cat allergens

XX and to diagnose sensitivity to allergens.

XX Claim 2: Fig 4; 70pp; English.

Poly-A mRNA from cat parotid and mandibular glands was used to
 CC produce cDNA clones for both chain 1 and chain 2 of TRFP. These
 CC clones were then used to screen a cat genomic library. Chain 1
 CC exists in two forms having different leader sequences (A and B).
 CC The sequence can be used to express the protein and peptide derivs.
 CC which stimulate T-cells in persons allergic to cats. The peptides
 CC can be used to reduce/eliminate the allergic response partic. by
 CC modification of lymphokine prodn. by the T-cells. They can also be
 CC used to identify epitopes responsible for sensitivity. The DNA can
 CC be used to detect comparable sequence in other species, and also

for prodn. of modified forms of TRFP esp. showing reduced binding to IgE and thus reduced tendency to cause adverse reactions. See also AAR12119-R12123.

Sequence 109 AA;

Query Match	90.2%;	Score 498;	DB 12;	Length 109;
Best Local Similarity	-92.8%;	Pred. No. 4.5e-51;		
Matches 103; Conservative	3;	Mismatches 3;	Indels 2;	Gaps 1;

```

1  DTMGALVIALVALVTOALGVKAETOPIDYDVEFVAVANGNLLDLSITVYNTPEERIA 60
1  dtmgalvialvalvtoalgvkaetopidydyvefvavangnllldlsitvyn tpeeria
1  dtmgalvialvltqalgykmaetcpitydvtfavangnellldlsitvtnatpeeria 600
1  dtmgalvialvltqalgykmaetcpitydvtfavangnellldlsitvtnatpeeria

61  MKKIDQCVENGILSRIDSLVYTTITSSSKDQMGAVNYEDKLTITLR 111
61  mkkidqcvengilsrldslvyttitssskdqmgavnyedkltitlr

61  mkkidqcvengilsrldslvny--talneycmgaavntvedkrlcttjcr 109
61  mkkidqcvengilsrldslvny--talneycmgaavntvedkrlcttjcr

```

Search completed: May 7, 2002, 12:13:28
db time: 200 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:16:04 ; Search time 12.5 Seconds
(without alignments)
199.829 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 111

Sequence: 1 DTKRGALLVALLVLTQALGV.....CMGEAVONTVEDLKLTLTGR 111

Scoring table: OLIGO

Searched: 212252 seqs, 22503292 residues

Word size: 0

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	111	1	US-07-662-193-4 Sequence 4, Appl
2	111	100.0	111	3	US-08-300-928C-6 Sequence 6, Appl
3	111	100.0	111	3	US-08-430-944D-6 Sequence 6, Appl
4	111	100.0	111	3	US-08-430-014-6 Sequence 6, Appl
5	111	100.0	111	3	US-08-431-184-6 Sequence 6, Appl
6	109	98.2	109	1	US-07-807-529A-6 Sequence 6, Appl
7	109	98.2	109	5	PCT-US93-02462-6 Sequence 6, Appl
8	92	82.9	92	3	US-08-300-928C-13 Sequence 13, Appl
9	92	82.9	92	3	US-08-430-944D-13 Sequence 13, Appl
10	92	82.9	92	3	US-08-430-014-13 Sequence 13, Appl
11	92	82.9	92	3	US-08-431-184-13 Sequence 13, Appl
12	83	74.8	109	1	US-07-662-193-5 Sequence 5, Appl
13	83	74.8	109	3	US-08-300-928C-8 Sequence 8, Appl
14	83	74.8	109	3	US-08-430-944D-8 Sequence 8, Appl
15	83	74.8	109	3	US-08-430-014-8 Sequence 8, Appl
16	83	74.8	109	3	US-08-431-184-8 Sequence 8, Appl
17	81	73.0	97	3	US-08-300-928C-10 Sequence 10, Appl
18	81	73.0	97	3	US-08-430-944D-10 Sequence 10, Appl
19	81	73.0	97	3	US-08-430-014-10 Sequence 10, Appl
20	81	73.0	97	3	US-08-431-184-10 Sequence 10, Appl
21	64	57.7	90	3	US-08-300-928C-14 Sequence 14, Appl
22	64	57.7	90	3	US-08-430-944D-14 Sequence 14, Appl
23	64	57.7	90	3	US-08-430-014-14 Sequence 14, Appl
24	64	57.7	90	3	US-08-431-184-14 Sequence 14, Appl
25	37	33.3	51	1	US-07-662-193-6 Sequence 6, Appl
26	37	33.3	51	3	US-08-300-928C-15 Sequence 15, Appl
27	37	33.3	51	3	US-08-430-944D-15 Sequence 15, Appl

28	37	33.3	51	3	US-08-430-014-15	Sequence 15, Appl
29	37	33.3	51	3	US-08-431-184-15	Sequence 15, Appl
30	36	32.4	83	1	US-07-662-193-7	Sequence 7, Appl
31	32	28.8	82	3	US-08-300-928C-16	Sequence 16, Appl
32	32	28.8	82	3	US-08-430-944D-16	Sequence 16, Appl
33	32	28.8	82	3	US-08-430-014-16	Sequence 16, Appl
34	32	28.8	82	3	US-08-431-184-16	Sequence 16, Appl
35	32	28.8	83	1	US-07-662-193-8	Sequence 8, Appl
36	26	23.4	26	1	US-07-807-529A-9	Sequence 9, Appl
37	26	23.4	26	3	US-08-300-928C-19	Sequence 19, Appl
38	26	23.4	26	3	US-08-430-944D-19	Sequence 19, Appl
39	26	23.4	26	3	US-08-430-014-19	Sequence 19, Appl
40	26	23.4	26	3	US-08-431-184-19	Sequence 19, Appl
41	26	23.4	26	5	PCT-US93-02462-9	Sequence 9, Appl
42	26	23.4	96	1	US-07-807-529A-39	Sequence 39, Appl
43	26	23.4	96	3	US-08-430-944D-103	Sequence 103, App
44	26	23.4	96	3	US-08-431-184-103	Sequence 103, App
45	23	20.7	27	3	US-08-300-928C-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-07-662-193-4
Sequence 4, Application US/07662193
Patent No. 5128991

GENERAL INFORMATION:

APPLICANT: Kuo, Mei-chang

APPLICANT: Bond, Julian

TITLE OF INVENTION: Improved Preparation of Cat Dander Allergens for Immunotherapeutic Purposes and Uses Thereof

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS: LAHIVE & COCKFIELD

ADDRESS: 60 State Street, Suite 510

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: APPLICATION NUMBER: US/07/662,193

FILING DATE: 19910228

CLASSIFICATION: 514

PRIOR APPLICATION DATA: APPLICATION NUMBER: US 07/662,276

FILING DATE: 28-FEB-1991

ATTORNEY/AGENT INFORMATION: NAME: Mandragouras, Amy E.

REGISTRATION NUMBER: P36,207

REFERENCE/DOCKET NUMBER: IML89-02AA/IPC-002CC/IMI-020

TELECOMMUNICATION INFORMATION: TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 4: SEQUENCE CHARACTERISTICS: LENGTH: 111 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: Internal

US-07-662-193-4

Query Match 100.0%; Score 111; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.1e-102;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Not in line w/ app

Wed May 8 09:01:30 2002

us-09-662-784-6.cli.ra1

Page 2

OY 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
DB 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
OY 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111
DB 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111

RESULT 2

US-08-300-928C-6
Sequence 6, Application US/08300928C
Patent No. 6019972

GENERAL INFORMATION:

APPLICANT: GEFTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991

ATTORNEY/AGENT INFORMATION:

NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002,605(IMT-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-300-928C-6

Query Match 100.0%; Score 111; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 1,1e-102; Indels 0; Gaps 0;
Matches 111; Conservative 0; Mismatches 0;

OY 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
DB 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
OY 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111
DB 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111

RESULT 3

US-08-430-944D-6
Sequence 6, Application US/08430944D
Patent No. 6025162

GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.

TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,944D
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)742-4214
TELEFAX: (617)227-7400

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-944D-6

Query Match 100.0%; Score 111; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 1,1e-102; Indels 0; Gaps 0;
Matches 111; Conservative 0; Mismatches 0;

OY 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
DB 1 DPMRGALVIALVLTQALGVKMAETCPIFYDVFPAVANGNELLDLSLTKVNAITEPERTA 60
OY 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111
DB 61 MKRIODCYVENGILSRVLDGLVMTTSSSKDKMGAVONTVEDLKLNTLGR 111

RESULT 4

US-08-430-014-6
Sequence 6, Application US/08430014
Patent No. 6048962

GENERAL INFORMATION:

APPLICANT: GEFTER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
TITLE OF INVENTION: PROTEIN (TRFP)
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESS: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/430,014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/300,928
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.60S(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-430-014-6

Query Match 100.0%; Score 111; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.1e-102;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTRGALVLAIVTQALGVKMAETCPITYDYFFAVANGNELLDLSLRKVANTEPERA 60
DB 1 DTRGALVLAIVTQALGVKMAETCPITYDYFFAVANGNELLDLSLRKVANTEPERA 60
QY 61 MKRIQCYVENGILSRVLDGLVMTTSSSKDCMGEAVONTVEDLKINTLGR 111
DB 61 MKRIQCYVENGILSRVLDGLVMTTSSSKDCMGEAVONTVEDLKINTLGR 111

RESULT 5
US-08-431-184-6
Sequence 6, Application US/08431184
Patent No. 6120769
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,184
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/430,014
FILING DATE: 27-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300,928
FILING DATE: 02-SEPT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-044DV3
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 111 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-431-184-6

Query Match 100.0%; Score 111; DB 3; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.1e-102;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTRGALVLAIVTQALGVKMAETCPITYDYFFAVANGNELLDLSLRKVANTEPERA 60
DB 1 DTRGALVLAIVTQALGVKMAETCPITYDYFFAVANGNELLDLSLRKVANTEPERA 60
QY 61 MKRIQCYVENGILSRVLDGLVMTTSSSKDCMGEAVONTVEDLKINTLGR 111
DB 61 MKRIQCYVENGILSRVLDGLVMTTSSSKDCMGEAVONTVEDLKINTLGR 111

RESULT 6
US-07-807-529A-6
Sequence 6, Application US/07807529A
Patent No. 5547669
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce L.
APPLICANT: Morgenstern, Jay
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Greenstein, Julia L.
APPLICANT: Kuo, Mei-chang
APPLICANT: Morville, Malcolm
TITLE OF INVENTION: RECOMBITOPE PEPTIDES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMLOGIC PHARMACEUTICAL CORPORATION
STREET: One Kendall Square, Building 600
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/807,529A
FILING DATE: 19911213
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
APPLICATION NUMBER: US 07/431,565
FILING DATE: 03-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Channing, Stacey L.
REGISTRATION NUMBER: 31,095
REFERENCE/DOCKET NUMBER: IPC-027/IMI-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 494-0060
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-807-529A-6

Wed May 8 09:01:30 2002

us-09-662-784-6.oli.ra1

Page 4

Query Match 98.2%; Score 109; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.1e-100;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVLAIVTQALGVKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK 62
DB 1 MRGALLVLAIVTQALGVKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK 60
QY 63 KIDCCYVENGSLISRVLDGLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 111
DB 61 KIDCCYVENGSLISRVLDGLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 109

RESULT 7
PCT-US93-02462-6
Sequence 6, Application PC/7059302462

GENERAL INFORMATION:
APPLICANT: Geffer, Malcolm L.
APPLICANT: Garman, Richard D.
APPLICANT: Greenstein, Julia L.
APPLICANT: Kuo, Wei-Chang
APPLICANT: Bliner, Thomas J.
APPLICANT: Morville, Malcolm
TITLE OF INVENTION: PEPTIDES USEFUL FOR TOLERIZATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/02462
FILING DATE: 19930325
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/006,116
FILING DATE: 15-JAN-1993
APPLICATION NUMBER: US 07/884,718
FILING DATE: 15-MAY-1992
APPLICATION NUMBER: 07/857,311
FILING DATE: 25-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-031PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-02462-6

Query Match 98.2%; Score 109; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.1e-100;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVLAIVTQALGVKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK 62
DB 1 MRGALLVLAIVTQALGVKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK 60
QY 63 KIDCCYVENGSLISRVLDGLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 111

DB 61 KIDCCYVENGSLISRVLDGLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 109

RESULT 8
US-08-300-928C-13
Sequence 13, Application US/08300928C
Patent No. 6019972

GENERAL INFORMATION:
APPLICANT: GEFFER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: WALTHAM
STATE: Massachusetts
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.6US(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-300-928C-13

Query Match 82.9%; Score 92; DB 3; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK KIDCCYVENGSLISRVLD 79
DB 1 VKMAETCPFYDVFPAVANGNELLDLSITKVNATEPERTAMK KIDCCYVENGSLISRVLD 60
QY 80 GLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 111
DB 61 GLVMTTSSSKDCMGEAVQNTVEDLKLNTLGR 92

RESULT 9

US-08-430-944D-13
Sequence 13, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.
TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:

```

; ADDRESSER: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430,944D
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/430,014
; FILING DATE: 27-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,928
; FILING DATE: 02-SEPT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-044DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-430-944D-13

Query Match      82.9%; Score 92; DB 3; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMAETCPIFYDYFFAVANGNELLDLSLTKVNATEPERTAMKIIDCYVENGSLRVID 79
DB 1 VKMAETCPIFYDYFFAVANGNELLDLSLTKVNATEPERTAMKIIDCYVENGSLRVID 60
QY 80 GLVMTTSSKDCMGFAVONTVEDLKLNTLGR 111
DB 61 GLVMTTSSKDCMGFAVONTVEDLKLNTLGR 92

RESULT 10
US-08-430-014-13
; Sequence 13, Application US/08430014
; Patent No. 6048962
; GENERAL INFORMATION:
; APPLICANT: GEFFER, Malcolm L. et al.
; TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
; TITLE OF INVENTION: PROTEIN (TRFP)
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430,014
```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/300,928
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: AMY E. MANDRAGOURAS
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: 002.60S(IMI-044)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-430-014-13

Query Match      82.9%; Score 92; DB 3; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMAETCPIFYDYFFAVANGNELLDLSLTKVNATEPERTAMKIIDCYVENGSLRVID 79
DB 1 VKMAETCPIFYDYFFAVANGNELLDLSLTKVNATEPERTAMKIIDCYVENGSLRVID 60
QY 80 GLVMTTSSKDCMGFAVONTVEDLKLNTLGR 111
DB 61 GLVMTTSSKDCMGFAVONTVEDLKLNTLGR 92

RESULT 11
US-08-431-184-13
; Sequence 13, Application US/08431184
; Patent No. 6120769
; GENERAL INFORMATION:
; APPLICANT: Bruce L. Rogers et al.
; TITLE OF INVENTION: A HUMAN T CELL REACTIVE FELINE PROTEIN
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,184
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/430,014
; FILING DATE: 27-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,928
; FILING DATE: 02-SEPT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: IMI-044DV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
```

Wed May 8 09:01:30 2002

us-09-662-784-6.011.ral

Page 6

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-431-184-13

Query Match 82.9%; Score 92; DB 3; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.1e-84;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VKMATECPIDYDFVFAVANGNELLDLSLTKVNATEPERTARKKIQDCYVENGSLISVLD 79
DB 1 VKMATECPIDYDFVFAVANGNELLDLSLTKVNATEPERTARKKIQDCYVENGSLISVLD 60
QY 80 GLVMTTSSSKDCMGRAVONTVEDIKNTLGR 111
DB 61 GLVMTTSSSKDCMGRAVONTVEDIKNTLGR 92

RESULT 12
US-07-662-193-5
Sequence 5, Application US/07662193
Patent No. 5328991
GENERAL INFORMATION:
APPLICANT: Kuo, Mei-chang
APPLICANT: Bond, Julian
TITLE OF INVENTION: Improved Preparation of Cat Dander
ALLERGENS FOR IMMUNOTHERAPEUTIC PURPOSES AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,193
FILING DATE: 19910228
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/662,276
FILING DATE: 28-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IML89-02AM/IPC-002CC/IMI-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-07-662-193-5

Query Match 74.8%; Score 83; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.8e-75;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTRGALLLVALLVYQALGYKMAETCPIDYDFVFAVANGNELLDLSLTKVNATEPERTA 60
DB 1 DTRGALLLVALLVYQALGYKMAETCPIDYDFVFAVANGNELLDLSLTKVNATEPERTA 60
QY 61 MKRIQDCYVENGSLISRVLDGLYM 83
DB 61 MKRIQDCYVENGSLISRVLDGLYM 83

RESULT 13
US-08-300-928C-8
Sequence 8, Application US/08300928C
Patent No. 6019972
GENERAL INFORMATION:
APPLICANT: GEFFER, Malcolm L. et al.
TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
STREET: 610 LINCOLN STREET
CITY: MALDEN
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02145
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/300,928C
FILING DATE: September 2, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,529
FILING DATE: December 13, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AMY E. MANDRAGOURAS
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 002.60US(IMI-044)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 109 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-300-928C-8

Query Match 74.8%; Score 83; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.8e-75;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTRGALLLVALLVYQALGYKMAETCPIDYDFVFAVANGNELLDLSLTKVNATEPERTA 60
DB 1 DTRGALLLVALLVYQALGYKMAETCPIDYDFVFAVANGNELLDLSLTKVNATEPERTA 60
QY 61 MKRIQDCYVENGSLISRVLDGLYM 83
DB 61 MKRIQDCYVENGSLISRVLDGLYM 83

RESULT 14
US-08-430-944D-8
Sequence 8, Application US/08430944D
Patent No. 6025162
GENERAL INFORMATION:
APPLICANT: Bruce L. Rogers et al.

```

1  TITLE OF INVENTION:  A HUMAN T CELL REACTIVE FELINE PROTEIN
2  TITLE OF INVENTION:
3  NUMBER OF SEQUENCES:  103
4  CORRESPONDENCE ADDRESSES:
5  ADDRESSEE:  LAHIVE & COCKFIELD, LLP
6  STREET:  28 State Street
7  CITY:  Boston
8  STATE:  Massachusetts
9  COUNTRY:  USA
10 ZIP:  02109
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  PatentIn Release #1.0, Version #1.25
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/08/430,944D
20 FILING DATE:  28-APR-1995
21
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER:  US 08/430,014
24 FILING DATE:  27-APR-1995
25
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER:  US 08/300,928
28 FILING DATE:  02-SEPT-1994
29
30 ATTORNEY/AGENT INFORMATION:
31 NAME:  Amy E. Mandragouras
32 REGISTRATION NUMBER:  36,207
33 REFERENCE/DOCKET NUMBER:  IMI-044DV2
34
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE:  (617)227-7400
37 TELEFAX:  (617)742-4214
38
39 INFORMATION FOR SEQ ID NO:  8:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH:  109 amino acids
42 TYPE:  amino acid
43 TOPOLOGY:  linear
44
45 MOLECULE TYPE:  protein
46
47 US-08-430-944D-8

```

```

Query Match                    74.8%; Score 83; DB 3; Length 109;
Best Local Similarity        100.0%; Pred. No. 5.8e-75;
Matches      83; Conservative    0; Mismatches    0; Indels    0; Gaps    0.

QY      1 DTMRCALLVLTALVTQAAGVKMAETCPITYDYFFAVANGNELLLDLSLTKVNATEPERTA 60
          |||||||
Db       1 DTMRCALLVLTALVTQAAGVKMAETCPITYDYFFAVANGNELLLDLSLTKVNATEPERTA 60
          |||||||

QY      61 MKKIODCYVENGLISRVLDTGVM 83
          |||||||
Db       61 MKKIODCYVENGLISRVLDTGVM 83
          |||||||

RESULT   15
US-08-430-014-8
; Sequence 8, Application US/08430014
; Patent No. 6048962
; GENERAL INFORMATION:
; APPLICANT: GETTER, Malcolm L. et al.
; TITLE OF INVENTION: PEPTIDES FOR HUMAN T CELL REACTIVE FELINE
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMUNOLOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

1 SOFTWARE: ASCII text
2 CURRENT APPLICATION DATA:
3 APPLICATION NUMBER: US/08/430,014
4
5 FILING DATE:
6 CLASSIFICATION:
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: 08/300,928
9
10 FILING DATE:
11 CLASSIFICATION:
12 ATTORNEY/AGENT INFORMATION:
13 NAME: AMY E. MANDRAGOURAS
14 REGISTRATION NUMBER: 36,247
15 REFERENCE/DOCKET NUMBER: 002.6US(IMT-044)
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (617) 227-7400
18 TELEFAX: (617) 227-5941
19 INFORMATION FOR SEQ ID NO: 8:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 109 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25
26 US-08-430-014-8

```

[illegible]

Search completed: May 7, 2002, 12:17:53
Job time: 109 sec

Wed May 8 09:01:30 2002

us-09-662-784-6.oli.ra1

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:17:39 ; Search time 22.16 Seconds
(without alignments)
732.682 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 111
Sequence: 1 DTMRGALVIALVLTQALGV.....CMGEAVONTVEKLKLTGLR 111

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 473505 seqs, 146272329 residues

Word size : 0

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	7.2	119	13	09YGD0
2	8	7.2	138	13	09PSW3
3	8	7.2	138	13	09PSW2
4	8	7.2	181	2	087052
5	8	7.2	266	2	09KW39
6	8	7.2	318	11	099NE4
7	8	7.2	322	4	09Y5Y7
8	8	7.2	322	4	09UNF4
9	8	7.2	390	2	09AEX7
10	8	7.2	390	2	09AEX5
11	8	7.2	401	5	09HT26
12	7	6.3	61	5	061029
13	7	6.3	73	5	061042
14	7	6.3	73	5	09XU39
15	7	6.3	74	5	061023
16	7	6.3	77	2	09KG69
17	7	6.3	78	2	026881
18	7	6.3	87	2	003539
19	7	6.3	88	5	061047

20	7	6.3	94	5	061030	061030 trypanosoma
21	7	6.3	102	5	061033	061033 trypanosoma
22	7	6.3	103	2	09RRB9	09rrb9 deinococcus
23	7	6.3	107	5	061050	061050 trypanosoma
24	7	6.3	112	5	09VLE8	09vle8 dirosophila
25	7	6.3	113	5	061052	061052 trypanosoma
26	7	6.3	115	2	09RD28	09rd28 streptomyces
27	7	6.3	115	5	061046	061046 trypanosoma
28	7	6.3	119	5	061034	061034 trypanosoma
29	7	6.3	122	5	015774	015774 trypanosoma
30	7	6.3	123	5	P90602	P90602 trypanosoma
31	7	6.3	123	5	015773	015773 trypanosoma
32	7	6.3	123	5	061027	061027 trypanosoma
33	7	6.3	125	5	061025	061025 trypanosoma
34	7	6.3	126	5	061021	061021 trypanosoma
35	7	6.3	126	5	061036	061036 trypanosoma
36	7	6.3	127	5	061035	061035 trypanosoma
37	7	6.3	128	5	P90603	P90603 trypanosoma
38	7	6.3	133	5	061022	061022 trypanosoma
39	7	6.3	136	5	015775	015775 trypanosoma
40	7	6.3	139	5	P90601	P90601 trypanosoma
41	7	6.3	139	5	061037	061037 trypanosoma
42	7	6.3	139	13	P79799	P79799 micrurus co
43	7	6.3	143	5	015776	015776 trypanosoma
44	7	6.3	147	5	061043	061043 trypanosoma
45	7	6.3	163	5	018679	018679 trypanosoma

ALIGNMENTS

RESULT 1
ID 09YGD0 PRELIMINARY; PRT: 119 AA.
AC 09YGD0;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE LAMPRLN L-0.9-10.
GN L-0.9.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_Taxid=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Matl A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RT "Characterization of lampirin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Robson P., Youson J.H., Keeley F.W.;
RL Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF097759; AAC97500.1; JOINED.
DR EMBL; AF097758; AAC97500.1; JOINED.
DR EMBL; AF097757; AAC97498.1; JOINED.
DR EMBL; AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 119 AA: 1145 MW: BB608EE263DF03E0 CRC64;

Query Match 7.2%; Score 8; DB 13; Length 119;

Best Local Similarity 100.0%; Pred. No. 3.6; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 ALLVALL 13
DB 7 ALLVALL 14

RESULT 2
ID Q9PSM2 PRELIMINARY; PRT: 138 AA.
AC Q9PSM2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE LAMPRIIN L-0.8-12.
GN L-0.8
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Slatetz E., Maiti A., Rawat M., Youson J.H.,
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Robson P., Youson J.H., Keelley F.W.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097756; AAC97499.1; JOINED.
SQ SEQUENCE 138 AA; 13094 MW; AB2F35D0D25D7580 CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 13; Length 138;
Matches 8; Conservative 100.0%; Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ALLVALL 13
Db 7 ALLVALL 14

RESULT 3
ID Q9PSM2 PRELIMINARY; PRT: 138 AA.
AC Q9PSM2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE LAMPRIIN L-0.9-12.
GN L-0.9
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Slatetz E., Maiti A., Rawat M., Youson J.H.,
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Robson P., Youson J.H., Keelley F.W.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF097759; AAC97501.1; JOINED.
DR EMBL: AF097758; AAC97501.1; JOINED.
SQ SEQUENCE 138 AA; 13144 MW; AB2F35D0C74924D0 CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 13; Length 138;
Matches 8; Conservative 100.0%; Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 6 ALLVALL 13
Db 7 ALLVALL 14
RESULT 4
ID Q87052 PRELIMINARY; PRT: 181 AA.
AC Q87052;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 248F PROTEIN (FRAGMENT).
GN 248F.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z17561;
RA Fallarino A.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ231108; CAI13150.1;
DR InterPro: IPR003764; NagaA.
DR Pfam: PF02612; NagaA; 1.
FT NON_TER 1
FT NON_TER 181
SQ SEQUENCE 181 AA; 19476 MW; 8527718DAC9DD199 CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 2; Length 181;
Matches 8; Conservative 100.0%; Pred. No. 5; Mismatches 0; Indels 0; Gaps 0;

Qy 95 EAVONTVE 102
Db 150 EAVONTVE 157

RESULT 5
ID Q9MW39 PRELIMINARY; PRT: 266 AA.
AC Q9MW39;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE COMPONENT OF TYPE IV SECRETION SYSTEM.
GN VYRBP.
OS Wolbachia sp. wkueto.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Wolbachieae; Wolbachia.
OX NCBI_TaxID=130485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WKUETO;
RA Masui S.;
RL "Air genes for the type IV secretion system of Wolbachia wkueto.";
DR Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB045235; BAA97440.1;
SQ SEQUENCE 266 AA; 31112 MW; 23C42E77A756B5E2 CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 2; Length 266;
Matches 8; Conservative 100.0%; Pred. No. 6; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LVIATLV 14
Db 7 LVIATLV 14

RESULT 6
Q99NE4 PRELIMINARY; PRT; 318 AA.
AC Q99NE4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYALURONAN RECEPTOR PRECURSOR.
GN LYVE-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-DIGESTIVE TRACT;
RX MEDLINE-99156989; PubMed-10037799;
RA Banerji S.;
RT "LYVE-1, a new homologue of the CD44 glycoprotein is a lymph-specific receptor for hyaluronan. LYVE-1, a new homologue of the CD44 glycoprotein is a lymph-specific receptor for hyaluronan.";
RL J. Cell Biol. 144:789-801(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-DIGESTIVE TRACT;
RA Prevost R., Banerji S., Ferguson D.J.P., Clasper S., Jackson D.G.;
RT "Mouse LYVE-1 is an endocytic receptor for hyaluronan in lymphatic endothelium.";
RL J. Biol. Chem. 0:0-0(0).
DR EMBL: AJ311501; CAC33082.1; -
FT SIGNAL 1 23
FT CHAIN 24 318 LYVE-1.
SQ SEQUENCE 318 AA; 34641 MW; 1248974A16113330 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALLVALL 13
Db 235 ALLVALL 242

RESULT 7
Q9Y5Y7 PRELIMINARY; PRT; 322 AA.
AC Q9Y5Y7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LYMPHATIC ENDOTHELIUM-SPECIFIC HYALURONAN RECEPTOR LYVE-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99156989; PubMed-10037799;
RA Banerji S., Ni J., Wang S.X., Clasper S., Su J., Tammi R., Jones M., Jackson D.G.;
RT "LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph-specific receptor for hyaluronan.";
RL J. Cell Biol. 144:789-801(1999).
DR EMBL: AF118108; AAD42764.1; -
DR HSSP: P98066; ITSG.
DR InterPro: IPR000538; Link.
DR Pfam: PF00193; Xlink: 1.
DR ProDom: PD000918; Link: 1.
DR SMART: SM00445; Link: 1.
RW Receptor.
SQ SEQUENCE 322 AA; 35238 MW; 0BBEA56729CEFF7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALLVALL 13
Db 239 ALLVALL 246

RESULT 8
Q9UNF4 PRELIMINARY; PRT; 322 AA.
AC Q9UNF4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HYALURONIC ACID RECEPTOR.
GN HAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Winkelman J.C., Basu S., Ozdemir E., Blough R.I.;
RT "HAR: a novel homolog of CD44 and putative hyaluronidic acid receptor encoded by a gene on human chromosome 11p15.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF127670; AAD49220.2; -
DR HSSP: P98066; ITSG.
DR InterPro: IPR000538; Link.
DR Pfam: PF00193; Xlink: 1.
DR PRINTS: PR01265; LINKMODULE.
DR ProDom: PD000918; Link: 1.
DR SMART: SM00445; Link: 1.
RW Receptor.
SQ SEQUENCE 322 AA; 35213 MW; 8B4D6D623F52D559 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALLVALL 13
Db 239 ALLVALL 246

RESULT 9
Q9AEX7 PRELIMINARY; PRT; 390 AA.
AC Q9AEX7;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE VARIABLE SURFACE PROTEIN VSPF.
GN VSPF.
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
OX NCBI_TaxID=159;
RN [1]
RP SEQUENCE FROM N.A.
RA McCaman M., Gabe J.;
RT "Serpulina hyodysenteriae (B204) contains eight gene copies related to a 39 kDa surface protein.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY027775; AK14801.1; -
SQ SEQUENCE 390 AA; 42836 MW; 9B20084CEB1A2B9E CRC64;

Query Match
7.2%; Score 8; DB 2; Length 390;

Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 94 GEVONTV 101
| | | | | | | | | |
DB 247 GEVONTV 254

RESULT 10

Q9AEK5 PRELIMINARY; PRT; 390 AA.
ID 09AEK5
AC 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE VARIABLE SURFACE PROTEIN VSPH.
GN VSPH.
OS Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
OX NCBI_TaxID=159;
RN [1]
RP SEQUENCE FROM N.A.
RT McManan M., Gabe J.;
FT a 39 kDa surface protein.;
RL Submitted (FBI-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL:AF02775; AAK14603.1;
SQ SEQUENCE 390 AA; 42816 MW; C3B635C9BC36E08 CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 2; Length 390;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 GEVONTV 101
| | | | | | | | | |
DB 245 GEVONTV 252

RESULT 11

Q9HT26 PRELIMINARY; PRT; 401 AA.
ID 09HT26
AC 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE PROBABLE MFS TRANSPORTER.
GN PASS48
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PAOI;
RX MEDLINE=20437373; PubMed=10984043;
RA Stoyer C.K., Pham X.-O.T., Elyan A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S., Hutnagle M.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltz L., Tolmach E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Conlister S., Poirier K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong S.K.S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock S.E.W., Lory S., Olson M.V.;
RT Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.
RL Nature 406:959-964 (2000).
DR EMBL:AE004967; AAC08933.1;
DR InterPro: IPR001958; TCR_TetB;
DR InterPro: IPR001411; TCR_TetB;
DR PRINTS: PR01035; TCRTEPA;
DR PRINTS: PR01036; TCRTEPB;
KW Complete proteome.
SQ SEQUENCE 401 AA; 42247 MW; D8D9EABDABD509E CRC64;

Query Match
Best Local Similarity 7.2%; Score 8; DB 2; Length 401;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LLEVALLV 14
| | | | | | | | | |
DB 281 LLEVALLV 288

RESULT 12

O61029 PRELIMINARY; PRT; 61 AA.
ID 061029
AC 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
DT 01-AUG-1998 (TREMblrel. 07, last annotation update)
DE MUCIN-LIKE PROTEIN (FRAGMENT).
GN EMUCE-21E22BIS.
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CL-BRENNER;
RA Di Nola J.M., D'Orso I., Aslund L., Sanchez D.O., Frasch A.C.C.;
RL J. Biol. Chem. 0:0-0(1998).
DR EMBL:AF036418; AAC14228.1;
FT NON-TER 61
SQ SEQUENCE 61 AA; 6192 MW; 95ABCTD84C05B587 CRC64;

Query Match
Best Local Similarity 6.3%; Score 7; DB 5; Length 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALLVIAL 12
| | | | | | | | | |
DB 10 ALLVIAL 16

RESULT 13

O61042 PRELIMINARY; PRT; 64 AA.
ID 061042
AC 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
DT 01-AUG-1998 (TREMblrel. 07, last annotation update)
DE MUCIN-LIKE PROTEIN (FRAGMENT).
GN EMUCE-13.
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CL-BRENNER;
RA Di Nola J.M., D'Orso I., Aslund L., Sanchez D.O., Frasch A.C.C.;
RL J. Biol. Chem. 0:0-0(1998).
DR EMBL:AF036439; AAC14242.1;
FT NON-TER 64
SQ SEQUENCE 64 AA; 6673 MW; A2B55B87D05608FF CRC64;

Query Match
Best Local Similarity 6.3%; Score 7; DB 5; Length 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ALLVIAL 12
| | | | | | | | | |
DB 14 ALLVIAL 20

RESULT 14

09XU39
ID 09XU39 PRELIMINARY; PRT; 73 AA.

AC 09XU39;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE F14F8.8 PROTEIN.
GN F14F8.8.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]

RN SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Smaison N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: 292782; CAB07183.1;
SQ SEQUENCE 73 AA; 7210 MW; 2F5F985045CB14ID CRC64;

Query Match 6.3%; Score 7; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LVLALL 13
Db 5 LVLALL 11

RESULT 15
O61023 PRELIMINARY; PRT; 74 AA.

AC 061023;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)
DE MUCIN-LIKE PROTEIN (FRAGMENT).
GN EMUCE-4.

OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5693;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL-BRENNER;
RA Di Nola J.M., D'Orso I., Aslund L., Sanchez D.O., Frasch A.C.C.;
RL J. Biol. Chem. 0:0-0(1998).
DR EMBL: AF036411; AAC14222.1;
FT NON_TER 74
SQ SEQUENCE 74 AA; 7743 MW; 734CC37663E21401 CRC64;

Db 14 ALLVIAL 20

Search completed: May 7, 2002, 12:20:52
Job time: 193 sec

Query Match 6.3%; Score 7; DB 5; Length 74;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 ALLVIAL 12
|||||||

Wed May 8 09:01:32 2002

us-09-662-784-6.oil.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:17:59 ; Search time 11.68 Seconds

(without alignments)
348.441 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 111
Sequence: 1 DFMRGALVIALVLTQALGV.....CMGEAVQNTVEDEKLNTLGR 111

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 100059 seqs, 36664827 residues

Word size: 0

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt.39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	98.2	109	1	FEL2_FELCA
2	8	7.2	139	1	LAMP_PETMA
3	8	7.2	299	1	YCXC_CVAPA
4	8	7.2	367	1	Y4HA_RHISN
5	8	7.2	378	1	NAGA_VIBCH
6	8	7.2	1106	1	ITR7_RAT
7	8	7.2	1179	1	ITR7_MOUSE
8	8	7.2	1181	1	ITR7_HUMAN
9	8	7.2	141	1	YEDD_SALTY
10	7	6.3	146	1	LYC2_PIG
11	7	6.3	148	1	LYC3_PIG
12	7	6.3	250	1	BACR_HALS4
13	7	6.3	322	1	QOX2_BACSU
14	7	6.3	332	1	CI82_CAVPO
15	7	6.3	337	1	YE71_HAEIN
16	7	6.3	397	1	UL50_HCMVA
17	7	6.3	400	1	ARGD_MYCTU
18	7	6.3	420	1	HISZ_SYNP7
19	7	6.3	424	1	ONB2_BACSU
20	7	6.3	434	1	GSA_CERY
21	7	6.3	531	1	RO60_DEIRA
22	7	6.3	574	1	ILVI_ECOLI
23	7	6.3	707	1	MM09_RABIT
24	7	6.3	712	1	MM09_BOVIN
25	7	6.3	729	1	CUL6_CAEEL
26	7	6.3	899	1	PR06_YEAST
27	7	6.3	1189	1	ITAH_HUMAN
28	7	6.3	1345	1	VCAP_HSV6U
29	7	6.3	3412	1	POLG_TBREVS
30	7	6.3	3414	1	POLG_TBREVA
31	7	6.3	3414	1	POLG_TBREVA
32	6	5.4	80	1	ITF_HUMAN
33	6	5.4	100	1	SECG_AOUAE

34	6	5.4	101	1	VNST_BUNYM	P16494 bunyamwera
35	6	5.4	101	1	VNST_MAGV	P16605 maguari vir
36	6	5.4	106	1	KACA_RAT	P01836 rattus norv
37	6	5.4	108	1	INS_BRARE	O73727 brachydanio
38	6	5.4	114	1	DCHS_LACBU	P04193 lactobacilli
39	6	5.4	115	1	MERT_SHEPU	O54462 shewanella
40	6	5.4	120	1	CHH4_PENMO	O97386 penaeus mon
41	6	5.4	123	1	YBAV_ECOLI	P77415 escherichia
42	6	5.4	125	1	RNP_SPAEH	P16418 spallax leuc
43	6	5.4	127	1	CHMO_BACSU	P19080 bacillus su
44	6	5.4	128	1	Y070_TREPA	O83109 treponema p
45	6	5.4	144	1	VE6_COPY	O89808 canine oral

ALIGNMENTS

RESULT	1	FEL2_FELCA	STANDARD:	PRT:	109 AA.
AC	P30440:				
DT	01-APR-1993 (Rel. 25, Created)				
DT	01-APR-1993 (Rel. 25, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	MAJOR ALLERGEN I POLYPEPTIDE CHAIN 2 PRECURSOR (ALLERGEN FEL D 1-B)				
DE	(FEL D 1-B) (ALLERGEN CAT-1) (AG4) (FDI).				
CN	CH2.				
OS	Felis silvestris catus (Cat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.				
OX	NCBI_TaxId=9685;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 18-100.				
RA	MEDLINE=92052157; PubMed=1946388;				
RA	Morgenstern J.P., Griffith I.J., Brauer A.W., Rogers B.L.,				
RA	Bond J.F., Chapman M.D., Kuo M.-C.;				
RT	"Amino acid sequence of Fel di, the major allergen of the domestic				
RT	cat: protein sequence analysis and cDNA cloning.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Liver;				
RC	MEDLINE=92241678; PubMed=1572548;				
RA	Griffith I.J., Craig S., Pollock J., Yu X.-B., Morgenstern J.P.,				
RA	Rogers B.L.;				
RT	"Expression and genomic structure of the genes encoding FdI, the				
RT	major allergen from the domestic cat.";				
RL	Gene 113:263-268(1992).				
RN	[3]				
RP	SEQUENCE OF 18-37, AND CHARACTERIZATION.				
RA	MEDLINE=91287714; PubMed=1712068;				
RA	Duffort O.A., Carreira J., Nicli G., Polo F., Lombardero M.;				
RT	"Studies on the biochemical structure of the major cat allergen Felis				
RT	domesticus I.";				
RL	Mol. Immunol. 28:301-309(1991).				
RN	[4]				
RP	CHARACTERIZATION.				
RA	MEDLINE=84265679; PubMed=6747135;				
RA	Letlerman K., Ohman J.L. Jr.;				
RT	"Cat allergen I: Biochemical, antigenic, and allergenic properties.";				
RL	J. Allergy Clin. Immunol. 74:147-153(1984).				
CC	-1- SUBUNIT: HETEROETRAMER COMPOSED OF TWO NON-COVALENTLY LINKED				
CC	DISULFIDE-LINKED HETERODIMER OF CHAINS 1 AND 2.				
CC	-1- ALTERNATIVE PRODUCTS: THE LONG (CH2L) AND THE SHORT (CHS) FORMS				
CC	MAY ARISE BY ALTERNATIVE SPLICING OR MAY REPRESENT DIFFERENT				
CC	ALLELES OF THE CH2 GENE. THE SEQUENCE SHOWN HERE IS THAT OF THE				
CC	LONG FORM (CH2L).				
CC	-1- TISSUE SPECIFICITY: THE LONG FORM IS PREFERENTIALLY EXPRESSED IN				
CC	THE SALIVARY GLAND, WHILE THE SHORT FORM IS PREFERENTIALLY				
CC	EXPRESSED IN THE SKIN.				
CC	-1- DISEASE: MAJOR ALLERGEN PRODUCED BY THE DOMESTIC CAT.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: W7341; AAC41616.1; -
 CC EMBL: X62478; AAC44345.1; -
 CC PIR: JC1127; JC1127.
 CC PIR: JC1145; JC1145.
 CC Allergen: glycoprotein; Signal; Polymorphism; Alternative splicing.
 CC SIGNAL 1 17
 CC CHAIN 18 109
 CC CARBOHYD 50 50
 CC VARSPLIC 82 82
 CC VARSPLIC 83 109
 CC VARSPLIC 72 72
 CC VARSPLIC 72 72
 CC VARSPLIC 74 75
 CC VARSPLIC 82 83
 CC VARSPLIC 85 85
 CC VARSPLIC 86 86
 CC VARSPLIC 86 86
 CC VARSPLIC 87 88
 CC VARSPLIC 89 89
 CC VARSPLIC 96 96
 CC VARSPLIC 105 105
 CC VARSPLIC 24 24
 CC VARSPLIC 32 32
 CC VARSPLIC 109 AA; 11854 MM; 857FB9CD76036CB9 CRC64;
 SO SEQUENCE

Query Match 98.2%; Score 109; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 1e-100;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 MRGALLVTLALVTQALGKMAKRCPIFYDVFPAVANGNELLDLSLTKNATPEPRMTAK 62
 1 MRGALLVTLALVTQALGKMAKRCPIFYDVFPAVANGNELLDLSLTKNATPEPRMTAK 60

Qy 63 KIDCVYENGSLSPVLDGIWTTSSSKDCEANGNONVDEKLTNR 111
 61 KIDCVYENGSLSPVLDGIWTTSSSKDCEANGNONVDEKLTNR 109

Db 61 KIDCVYENGSLSPVLDGIWTTSSSKDCEANGNONVDEKLTNR 109

RESULT 2
 LAMP_PETMA STANDARD; PRT: 139 AA.
 ID LAMP_PETMA
 AC P33575; P33576; (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 30-MAY-2000 (Rel. 39, last annotation update)
 DE LAMPRIIN 0.9 PRECURSOR (CARTILAGE MATRIX PROTEIN).
 OS Petromyzon marinus (Sea Lamprey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Petromyzon.
 NCBI_Taxid=7757;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-44.
 RC TISSUE=cartilage;
 RX MEDLINE=93123269; PubMed=7678258;
 RA Robson P., Wright G.M., Stoltz E., Maitl A., Rawat M., Youson J.H.,
 RA Kealey F.W.;
 RA Characterization of lampriin, an unusual matrix protein from lamprey
 RT cartilage. Implications for evolution, structure, and assembly of
 RT elastin and other fibrillar proteins.;
 RL J. Biol. Chem. 268:1440-1447 (1993).
 CC -1- FUNCTION: SELF-AGGREGATING PROTEIN THAT MAKES PART OF THE SOLUBLE
 CC FORM OF LAMPRIIN.
 CC -1- SUBUNIT: THE POLYMERIC LAMPRIIN CHAINS SELF-AGGREGATE TO FORM
 CC FIBERS AND HAVE SECONDARY STRUCTURES PARTICULARLY RICH IN BETA-
 CC SHEETS AND IN BETA-TURNS.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS: 0.9-12 (SHOWN HERE) AND 0.9-10;
 CC SEEM TO BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L05925; AAA49269.1; -
 CC EMBL: L05924; AAA49268.1; -
 CC PIR: C45051; C45051.
 CC PIR: A45051; A45051.
 CC Cartilage; Repeat; Connective tissue; Signal; Alternative splicing.
 CC SIGNAL 1 19
 CC CHAIN 20 139
 CC DOMAIN 42 46
 CC REPEAT 42 46
 CC REPEAT 47 51
 CC REPEAT 52 56
 CC REPEAT 57 61
 CC REPEAT 62 66
 CC REPEAT 67 71
 CC REPEAT 72 76
 CC REPEAT 77 81
 CC REPEAT 82 86
 CC REPEAT 87 91
 CC REPEAT 92 96
 CC REPEAT 97 101
 CC REPEAT 102 106
 CC VARSPLIC 86 104
 CC VARSPLIC 139 AA; 13257 MM; E24BABB7A6CEC7C3 CRC64;
 SO SEQUENCE

Query Match 7.2%; Score 8; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 ALLVIAL 13
 7 ALLVIAL 14

Qy 6 ALLVIAL 13
 7 ALLVIAL 14

Db 7 ALLVIAL 14

RESULT 3
 YCXC_CYPARA STANDARD; PRT: 299 AA.
 ID YCXC_CYPARA
 AC P31606; P31607; (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 33, last sequence update)
 DT 01-FEB-1996 (Rel. 33, last sequence update)
 DT 20-AUG-2001 (Rel. 40, last annotation update)
 DE HYPOTHETICAL 32.8 KDA PROTEIN IN YCFC23-ABCF INTERGENIC REGION
 DE (ORF299).
 OS Cyanophora paradoxa.
 OC Cyanophyta; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
 OC Eukaryota.
 NCBI_Taxid=2762;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L8555 / PRINGSHEIM;
 RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
 RA Bryant D.A.;
 RA "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa";
 RT Plant Mol. Biol. Rep. 13:327-332(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L8555 / PRINGSHEIM;
 RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
 RA Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
 RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
 RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
 RT the genetic complexity of a primitive plastid.";
 RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
 RL Schwemmer W. (eds.);
 RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
 RL (1997).

[3]
RN SEQUENCE OF 138-299 FROM N.A.
RX MEDLINE=92201692; PubMed=151590;
RA Rhel E., Sturewalt V.L., Gasparich G.E., Bryant D.A.;
RT "The pasc genes of *Synechococcus* sp. FCC7002 and *Cyanophora paradoxa*:
RL Cloning and sequence analysis.";
RL Gene 112:123-128(1992).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U30821; AAA81303.1; -
CC EMBL: M86239; AAA65471.1; -
CC PIR: PS0372; PS0372.
DR InterPro: IPR001626; ABC-3.
DR Pfam: PF00950; ABC-3; 1.
KW Hypothetical protein; Cyanelle; Transmembrane.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 36 56 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 241 261 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
SQ SEQUENCE 299 AA: 32817 MW: 76830A20753CE66 CRC64;

Query Match 7.2%; Score 8; DB 1; Length 299;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 LVALLVLT 15
| | | | | | | |
DB 220 LVALLVLT 227

RESULT 4
Y4HA_RHISN STANDARD: PRT; 367 AA.
ID Y4HA_RHISN
AC P55471;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PUTATIVE IONIC TRANSPORTER Y4HA.
GN Y4HA.
OS *Rhizobium* sp. (strain NGR234).
OC Plasmid sym pNGR234.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freilberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between *Rhizobium* and legumes.";
RL Nature 387:394-401(1997).
CC -1- FUNCTION: POSSIBLE CATION TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: STRONG, TO E COLI CALCIUM/PROTON ANTIORTER (CHAA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: AE000075; AAB91689.1; -
KW Hypothetical protein; Transport; Ion transport; Transmembrane;
RN Plasmid.
FT TRANSMEM 12 32 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 221 241 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT TRANSMEM 347 367 POTENTIAL.
SQ SEQUENCE 367 AA: 38078 MW: 3AC0B829F2EFA2CB CRC64;

Query Match 7.2%; Score 8; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 GALLVLTAL 12
| | | | | | | |
DB 224 GALLVLTAL 231

RESULT 5
NAGA_VIBCH STANDARD: PRT; 378 AA.
ID NAGA_VIBCH
AC O32445; Q9KTA9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE N-ACETYLGLUCOSAMINE-6-PHOSPHATE DIACETYLASE (EC 3.5.1.25) (GLCNAC 6-P
DE DIACETYLASE).
GN NAGA OR VC0994.
OS *Vibrio* cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NON-O1 / 1148A;
RX MEDLINE=97446530; PubMed=9301118;
RA Yamano N., Oura N., Wang J., Fujishima S.;
RT "Cloning and sequencing of the genes for N-acetylglucosamine use that
RT non-O1.";
RL Biosci. Biotechnol. Blochem. 61:1349-1353(1997).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: N-ACETYL-D-GLUCOSAMINE 6-PHOSPHATE + H(2)O
CC = D-GLUCOSAMINE 6-PHOSPHATE + ACETATE.
CC -1- PATHWAY: N-ACETYL GLUCOSAMINE UTILIZATION PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE NAGA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D87820; BAA22834.1; -
 CC EMBL: AE004181; AAF94135.1; -
 CC TIGR: VC0994; -
 CC InterPro: IPR003764; Naga.1.
 CC Pfam: PF02612; Naga.1.
 CC Hydroxylase: Carbohydrate metabolism: Complete proteome.
 CC CONFLICT 79 79 T -> I (IN REF. 1).
 CC SEQUENCE 378 AA: 40956 MW: 34906344AA3F92A0F CRC64;

Query Match 7.28; Score 8; DB 1; Length 378;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 95 EAYONTVE 102
 DB 312 EAYONTVE 319

RESULT 6
 ID ITA7_RAT STANDARD; PRT: 1106 AA.
 AC Q63258; Q63026; Q63027;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INTEGRIN ALPHA-7 (H36-ALPHA7).
 GN ITGA7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X1B).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=92242309; PubMed=135319;
 RA Song W.K., Wang W., Foster R.F., Biesler D.A., Kaufman S.J.;
 RT "H36-alpha 7 is a novel integrin alpha chain that is developmentally
 RT regulated during skeletal myogenesis.";
 RL J. Cell Biol. 117:643-657(1992).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X1A AND ALPHA-7X1C).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=94171924; PubMed=8126096;
 RA Song W.K., Wang W., Sato H., Biesler D.A., Kaufman S.J.;
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
 RT muscle development: alternate forms, conformational change, and
 RT homologues with serine/threonine kinases and tyrosine phosphatases";
 RL J. Cell Sci. 106:1139-1152(1993).
 RN [3]
 RN TISSUE SPECIFICITY.
 RP MEDLINE=96197133; PubMed=8626012;
 RA Martin P.T., Kaufman S.J., Ktamer R.H., Sanes J.R.;
 RT "Synaptic integrins in developing, adult, and mutant muscle: selective
 RT association of alpha1, alphaA, and alphaB integrins with the
 RT neuromuscular junction";
 RL Dev. Biol. 174:125-139(1996).
 CC -1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 CC ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 CC DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 CC OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 CC SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF
 CC THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,
 CC VIABILITY AND FUNCTIONAL INTEGRITY.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A

CC DISRUPTIVE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: ALPHA-7X1A, ALPHA-7X1B
 CC (SIKON HERE) AND ALPHA-7X1C, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLE.
 CC EXPRESSED IN REPLICATING MYOBLASTS. IN DIFFERENTIATED MUSCLE
 CC FIBERS LOCALIZES BETWEEN FIBERS AND THE SURROUNDING MATRIX.
 CC ISOFORMS A AND B ARE EXPRESSED AT MYOTENDINOUS AND NEUROSCULAR
 CC JUNCTIONS; ISOFORM C IS EXPRESSED AT NEUROSCULAR JUNCTIONS AND
 CC AT EXTRASYNAPTIC SITES.
 CC -1- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING
 CC THE FORMATION OF SKELETAL MUSCLE. ISOFORMS A AND C ARE INDUCED
 CC UPON TERMINAL MYOGENIC DIFFERENTIATION; ISOFORM B IS PRESENT
 CC EARLIER IN REPLICATING CELLS AND DIMINISHES UPON DIFFERENTIATION.
 CC -1- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES IN THE EXTRACELLULAR
 CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X65036; CAA46170.1; -
 CC EMBL: X74293; CAA52346.1; -
 CC EMBL: X74294; CAA52347.1; -
 CC HSSP: P11215; IABX.
 CC InterPro: IPR000413; Integrin_alpha.
 CC Pfam: PF01839; FG-GAP. 5.
 CC Pfam: PF00357; Integrin_A; 1.
 CC PRINTS: PRO1185; INTEGRINA.
 CC SMART: SM00191; Int.alpha. 5.
 CC PROSITE: PS00242; INTEGRIN ALPHA: 1.
 CC Integrin: Cell adhesion: Receptor: Glycoprotein; Transmembrane;
 CC Extracellular matrix; Cytoskeleton; Repeat;
 CC Alternative splicing; ADP-ribosylation; Repeat;
 CC CHAIN 1 882
 FT CHAIN 1 882
 FT DOMAIN 886 1106
 FT DOMAIN 1004 1029
 FT DOMAIN 1030 1106
 FT REPEAT 17 55
 FT REPEAT 90 122
 FT REPEAT 162 194
 FT REPEAT 233 269
 FT REPEAT 290 328
 FT REPEAT 351 387
 FT REPEAT 409 449
 FT CA_BIND 300 308
 FT CA_BIND 362 370
 FT CA_BIND 420 428
 FT DOMAIN 885 885
 FT SITE 1032 1036
 FT DOMAIN 1082 1101
 FT REPEAT 1082 1085
 FT REPEAT 1090 1093
 FT REPEAT 1096 1101
 FT DISULFD 61 71
 FT DISULFD 108 131
 FT DISULFD 152 165
 FT DISULFD 467 474
 FT DISULFD 480 543
 FT DISULFD 609 615
 FT DISULFD 708 719
 FT DISULFD 866 920
 FT DISULFD 926 931
 FT CARBOHD 713 713
 FT CARBOHD 915 915
 FT CARBOHD 950 950
 CC -----
 CC 1. INTEGRIN ALPHA-7 HEAVY CHAIN (POTENTIAL).
 CC EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC FG-GAP 1.
 CC FG-GAP 2.
 CC FG-GAP 3.
 CC FG-GAP 4.
 CC FG-GAP 5.
 CC FG-GAP 6.
 CC FG-GAP 7.
 CC POTENTIAL.
 CC POTENTIAL.
 CC POTENTIAL.
 CC POLY-ARG.
 CC GFERR MOTIF.
 CC 3 X 4 AA REPEATS OF D-X-H-P.
 CC 1.
 CC 2.
 CC 3. BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC INTERCHAIN (BY SIMILARITY).
 CC BY SIMILARITY.
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 970 970 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1031 1106 LGFFRAKHPATVPOYHAKVILREDOQKEEKTIGORS
 FT VARSPLIC 1036 1106 MNGNSOMGSDAHPILADMPHILGPDGHPVSVA -> CG
 FT VARSPLIC 1036 1106 FFRNRPSSPPANRRAHILAVOPSPAMAGGPGTGMSSSS
 FT VARSPLIC 1036 1106 GRSTLRPIYPTQ (IN ISOFORM ALPHA-7X1A).
 FT VARSPLIC 1036 1106 RAKHPEATVPYHAKVILREDOQKEEKTIGORSNNGNS
 FT VARSPLIC 1036 1106 QMGSDAHPILADMPHILGPDGHPVSVA -> CAVPAOR
 FT VARSPLIC 1036 1106 ILSIY (IN ISOFORM ALPHA-7X1C).
 SO SEQUENCE 1106 AA; 121101 MW; 21B2A187837E01F6 CRC64;
 Query Match 7.2%; Score 8; DB 1; Length 1106;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 LLYVALLV 14
 Db 1018 LLYVALLV 1025
 RESULT 7
 ITA7_MOUSE STANDARD: PRT: 1179 AA.
 ID ITA7_MOUSE PRT: 1179 AA.
 AC Q61738; P70350; O88732; Q61737; O88731; Q61741;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE INTEGRIN ALPHA-7 PRECURSOR.
 GN ITGA7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Murinae; Mus.
 ON NCBI_Taxid=10090;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS ALPHA-7A; ALPHA-7X2B AND ALPHA-7X1X2).
 RC STRAIN=BAIB/C; TISSUE=Heart;
 RX MEDLINE=94075378; PubMed=8253814;
 RA Ziber B.L., Vu M.P., Waleh N., Crawford J., Ian C.-S., Kramer R.H.;
 RT "Alternative extracellular and cytoplasmic domains of the integrin
 alpha 7 subunit are differentially expressed during development.";
 RL J. Biol. Chem. 268:26773-26783(1993).
 RN SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X1X2A AND ALPHA-7X1X2B).
 RC STRAIN=129/SV;
 RA Saher G., Echtermeier F., Beier D.R., Poeschl E., Mayer U.;
 RT "Genomic organization and chromosomal localization of the mouse
 integrin alpha7 gene";
 RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
 RN SEQUENCE OF 1-70 FROM N.A.
 RC STRAIN=C57BL/6 x CBA;
 RX MEDLINE=96394366; PubMed=8798472;
 RA Ziber B.L., Kramer R.H.;
 RT "Identification and characterization of the cell type-specific and
 developmentally regulated alpha7 integrin gene promoter.";
 RL J. Biol. Chem. 271:22915-22922(1996).
 RN SEQUENCE OF 34-58.
 RC TISSUE=melanoma;
 RX MEDLINE=92198982; PubMed=1839357;
 RA Krimer R.H., Vu M.P., Cheng Y.F., Ramos D.M., Timpl R., Waleh N.;
 RT "Laminin-binding integrin alpha 7 beta 1: functional characterization
 and expression in normal and malignant melanocytes";
 RL Cell Regul. 2:805-817(1991).
 RN PARTIAL SEQUENCE FROM N.A. (ISOFORMS ALPHA-7A AND ALPHA-7B).
 RC STRAIN=C57BL/6 x BALB/C;
 RX MEDLINE=93366824; PubMed=8360188;
 RA Collo G., Starr L., Quaranta V.;
 RT "A new isoform of the laminin receptor integrin alpha 7 beta 1 is
 developmentally regulated in skeletal muscle.";
 RL J. Biol. Chem. 268:19019-19024(1993).
 RN [6]

RP TISSUE SPECIFICITY.
 RX MEDLINE=96197133; PubMed=8626012;
 RA Martin P.T., Kauffman S.J., Krimer R.H., Sanes J.R.;
 RT "Synaptic integrins in developing, adult, and mutant muscle: selective
 association of alpha1, alpha7A, and alpha7B integrins with the
 neuromuscular junction";
 RL Dev. Biol. 174:125-139(1996).
 RN [7]
 RN FUNCTION.
 RX MEDLINE=98016417; PubMed=9354797;
 RA Mayer U., Saher G., Fässler R., Bornemann A., Echtermeier F.,
 RA von der Mark H., Miosge N., Poeschl E., von der Mark K.;
 RT "Absence of integrin alpha 7 causes a novel form of muscular
 dystrophy";
 RL Nat. Genet. 17:318-323(1997).
 RN [8]
 RN ADP-RIBOSYLATION.
 RX MEDLINE=95238432; PubMed=7721841;
 RA Zolkiewska A., Moss J.;
 RT "Processing of ADP-ribosylated integrin alpha 7 in skeletal muscle
 myotubes";
 RL J. Biol. Chem. 270:9227-9233(1995).
 CC -1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF
 THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,
 VIABILITY AND FUNCTIONAL INTEGRITY. MICE CARRYING A ITGA7 NDL
 ALLELE ARE VIABLE AND FERTILE, BUT SHOW PROGRESSIVE MUSCULAR
 DYSTROPHY STARTING SOON AFTER BIRTH, BUT WITH A DISTINCT
 VARIABILITY IN DIFFERENT MUSCLE TYPES.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
 DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: ALPHA-7X1A, ALPHA-7X1B,
 ALPHA-7X2A, ALPHA-7X2B, ALPHA-7X1X2A AND ALPHA-7X1X2B (SHOWN
 HERE). MAY BE PRODUCED BY ALTERNATIVE SPLICING. THERE IS A
 COMBINATION OF AT LEAST FOUR ALTERNATIVE SPLICED DOMAINS, TWO
 EXTRACELLULAR (X1 AND X2) AND TWO CYTOPLASMIC (A AND B). A THIRD
 POTENTIAL ALTERNATIVE SPLICED CYTOPLASMIC DOMAIN (C) DOESN'T
 APPEAR TO BE EXPRESSED. SO FAR DETECTED ARE ISOFORMS ALPHA-7X1A,
 ALPHA-7X1B AND ALPHA-7X2B.
 CC -1- TISSUE SPECIFICITY: ISOFORMS CONTAINING SEGMENT X2 ARE FOUND IN
 ADULT HEART, LUNG AND SKELETAL MUSCLE. ISOFORMS CONTAINING SEGMENT
 X1 ARE EXPRESSED IN ADULT HEART, LUNG AND IN PROLIFERATING
 SKELETAL MYOBLASTS BUT NOT IN ADULT SKELETAL MUSCLE. ISOFORMS
 CONTAINING SEGMENT A ARE EXCLUSIVELY FOUND IN SKELETAL MUSCLE.
 CC ISOFORMS CONTAINING SEGMENT B ARE WIDELY EXPRESSED. IN MUSCLE
 FIBERS ISOFORMS CONTAINING SEGMENT A AND B ARE EXPRESSED AT
 MYOTENDINOUS AND NEUROMUSCULAR JUNCTIONS; ISOFORMS CONTAINING
 SEGMENT C ARE EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND AT
 EXTRASYNAPTIC SITES.
 CC -1- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING
 FORMATION OF SKELETAL MUSCLE. UNDIFFERENTIATED (REPLICATING)
 MYOBLASTS EXPRESS ISOFORMS CONTAINING SEGMENT B ONLY, WHEREAS
 DIFFERENTIATED MYOBLASTS EXPRESS ISOFORMS CONTAINING SEGMENTS A OR
 B.
 CC -1- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES OF THE EXTRACELLULAR
 DOMAIN IN SKELETAL MYOTUBES (IN VITRO).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: L23423; AAA16600.1; -

DR	EMBL:	U60419;	AAC52772.1;	-	-
DR	EMBL:	Y12380;	CAG73023.1;	-	JOINED.
DR	EMBL:	Y12383;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12384;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12385;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12386;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12387;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12388;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12389;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12390;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12391;	CAG73023.1;	JOINED.	
DR	EMBL:	Y12392;	CAG73023.1;	JOINED.	
DR	EMBL:	L23442;	AAA16599.1;	-	
DR	EMBL:	L12360;	CAA73024.1;	-	
DR	EMBL:	Y12363;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12364;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12365;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12386;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12387;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12388;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12389;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12390;	CAA73024.1;	JOINED.	
DR	EMBL:	Y12391;	CAA73024.1;	JOINED.	
DR	EMBL:	L23421;	AAA16598.1;	-	
DR	HSSP:	P11215;	IABX.		
DR	MGI:	MGI:102700;	Itga7.		
DR	InterPro:	IPR000413;	Integrin_alpha.		
DR	Pfam:	PF01839;	FG-GAP; 5.		
DR	Pfam:	PF00357;	Integrin_A_2.		
DR	PRINTS:	PR01185;	INTEGRINA.		
DR	SMART:	SMO0191;	Int_alpha; 5.		
DR	PROSITE:	PS00242;	INTEGRIN_ALPHA; 1.		
KW	Integral; Cell adhesion; Receptor; Glycoprotein; Transmembrane;				
KM	Signal; Extracellular matrix; Cytoskeleton; Repeat;				
KV	Alternative splicing; ADP-ribosylation; Calcium.				
FT	SIGNAL	1	33		
FT	CHAIN	34	1179		INTEGRIN ALPHA-7.
FT	CHAIN	34	955		INTEGRIN ALPHA-7 HEAVY CHAIN (POTE
FT	CHAIN	959	1179		INTEGRIN ALPHA-7 LIGHT CHAIN (POTE
FT	DOMAIN	34	1076		EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1077	1102		POTENTIAL.
FT	DOMAIN	1103	1179		CYTOSOLASTIC (POTENTIAL).
FT	REPEAT	50	87		FG-GAP 1.
FT	REPEAT	132	134		FG-GAP 2.
FT	REPEAT	154	226		FG-GAP 3.
FT	REPEAT	305	341		FG-GAP 4.
FT	REPEAT	362	400		FG-GAP 5.
FT	REPEAT	423	459		FG-GAP 6.
FT	REPEAT	481	521		FG-GAP 7.
FT	CA_BIND	372	380		POTENTIAL.
FT	CA_BIND	434	442		POTENTIAL.
FT	CA_BIND	492	500		POTENTIAL.
FT	DOMAIN	953	958		POLY-ARG.
FT	SITE	1105	1109		GFPKR MOTIF.
FT	DOMAIN	1155	1174		3 X 4 AA REPEATS OF D-X-H-D.
FT	REPEAT	1155	1158		1.
FT	REPEAT	1163	1166		2.
FT	REPEAT	1171	1174		3.
FT	DISULFID	94	103		BY SIMILARITY.
FT	DISULFID	140	163		BY SIMILARITY.
FT	DISULFID	184	197		BY SIMILARITY.
FT	DISULFID	539	546		BY SIMILARITY.
FT	DISULFID	552	615		BY SIMILARITY.
FT	DISULFID	681	687		BY SIMILARITY.
FT	DISULFID	781	792		BY SIMILARITY.
FT	DISULFID	939	993		INTERCHAIN (BY SIMILARITY).
FT	DISULFID	999	1004		BY SIMILARITY.
FT	CARBONYD	86	86		N-LINKED (GLCNAc . . .) (POTENTIAL)
FT	CARBONYD	784	784		N-LINKED (GLCNAc . . .) (POTENTIAL)
FT	CARBONYD	988	988		N-LINKED (GLCNAc . . .) (POTENTIAL)
FT	CARBONYD	1023	1023		N-LINKED (GLCNAc . . .) (POTENTIAL)
FT	CARBONYD	1043	1043		N-LINKED (GLCNAc . . .) (POTENTIAL)

FT	VARSPPLIC	224	267	MISSING (IN ISOFORM ALPHA-7X2A AND ISOFORM ALPHA-7X2B)
FT	VARSPPLIC	268	307	MISSING (IN ISOFORM ALPHA-7X1A AND ISOFORM ALPHA-7X2B)
Query Match	Best Local Similarity	7.2%	Score 8:	DB 1: Length 1179;
Matches	8: Conservative	0:	Mismatches	0: Indels 0: Gaps 0:
QY	7	LEVLAIVL 14		
DB	1091	LEVLAIVL 1098		
RESULT	8			
ITAT7_HUMAN	STANDARD:	PRT: 1181 AA.		
ID	Q13683; Q043197; Q09UEV2; Q09UETO; Q09NY89;			
AC	Q13683; Q043197; Q09UEV2; Q09UETO; Q09NY89;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	INTEGRIN ALPHA-7 PRECURSOR.			
GN	ITGA7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_Taxid	9606;			
[1]				
RP	SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X2B AND ALPHA-7X2DB).			
RC	TISSUE=retal heart, and Osteoblast;			
RX	MEDLINE=98139911; PubMed=9473524;			
RA	Leung E., Lim S.P., Berg R., Yang Y., Ni J., Wang S.-X.,			
RA	Krisanssen G.W.;			
RT	"A novel extracellular domain variant of the human integrin alpha 7			
RT	subunit generated by alternative intron splicing."			
RL	Biochem. Biophys. Res. Commun. 243:317-325(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).			
RX	MEDLINE=98250181; PubMed=9590299;			
RA	Hayashi Y.K., Chou F.-L., Engvall E., Ogawa M., Matsuda C.,			
RA	Hirabayashi S., Yokochi K., Zlobor B.L., Kramer R.H., Kaufman S.J.,			
RA	Ofman E., Goto T.-I., Nonaka I., Tsukahara T., Wang J.Z.,			
RA	Hofman E.P., Atahala K.;			
RT	Mutations in the integrin alpha7 gene cause congenital myopathy."			
RL	Nat. Genet. 19:94-97(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).			
RA	Vizirianakis I.S., Zlobor B.L., Kramer R.H.;			
RT	"Cloning of human integrin alpha-7 cDNA."			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.			
RC	TISSUE=Skeletal muscle;			
RX	MEDLINE=99333684; PubMed=10403775;			
RA	Vignier N., Moghadassadeh B., Gary F., Beckmann J., Mayer U.,			
RA	Guicheney P.;			
RT	"Structure, genetic localization, and identification of the cardiac			
RT	and skeletal muscle transcripts of the human integrin alpha7 gene			
RT	(TTGA7)."			
RL	Biochem. Biophys. Res. Commun. 260:357-364(1999).			
RN	[5]			
RP	SEQUENCE OF 34-45.			
RC	TISSUE=melanoma;			
RX	MEDLINE=92198982; PubMed=1839357;			
RA	Kramer R.H., Vu M.P., Cheng Y.F.,			
RA	Ramos D.M., Timpl R., Waleh N.;			
RT	"Laminin-binding integrin alpha beta 1: functional characterization			
RT	and expression in normal and malignant melanocytes."			
RL	Cell Regul. 2:805-817(1991).			
RN	[6]			
RP	PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2).			
RC	TISSUE=heart;			
RX	MEDLINE=94073378; PubMed=8253814;			
RA	Zlobor B.L., Vu M.P., Waleh N., Crawford J., Lin C.-S., Kramer R.H.;			
RT	"Alternative extracellular and cytoplasmic domains of the integrin			

RT alpha 7 subunit are differentially expressed during development.";
RL J. Biol. Chem. 268:26773-26783(1993).
[7]
RP SEQUENCE OF 1105-1181 FROM N.A. (ISOFORM ALPHA-7B).
RC TISSUE=Fetal muscle;
RX MEDLINE=94171924; PubMed=8126096;
RA Song W.K., Wang W., Sato H., Hiesler D.A., Kaufman S.J.;
RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
muscle development: alternate forms, conformational change, and
homologies with serine/threonine kinases and tyrosine phosphatases.";
RL J. Cell Sci. 106:1139-1152(1993).
[8]
RN PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7A).
RP TISSUE=Skeletal muscle;
RX MEDLINE=98012902; PubMed=9352853;
RA Basora N., Vachon P.H., Herring-Gillam F.E., Perrault N.,
RT Beaulieu J.-F.;
RT "Relation between integrin alpha7/beta1 expression in human intestinal
cells and enterocytic differentiation.";
RL Gastroenterology 113:1510-1521(1997).
[9]
RN FUNCTION.
RX MEDLINE=97453229; PubMed=9307969;
RA Zlober B.L., Chen Y.Q., Kramer R.H.;
RT "The laminin-binding activity of the alpha 7 integrin receptor is
defined by developmentally regulated splicing in the extracellular
domain.";
RL Mol. Cell Biol. 17:1723-1734(1997).
[10]
RN FUNCTION.
RX MEDLINE=20160722; PubMed=10694445;
RA Schoeber S., Mielenz D., Echtermeier F., Hapke S., Poeschl E.,
RT von der Mark H., Moch H., von der Mark K.;
RT "The role of extracellular and cytoplasmic splice domains of
alpha7-integrin in cell adhesion and migration on laminins.";
RL Exp. Cell Res. 255:303-313(2000).
[11]
RN TISSUE SPECIFICITY.
RP MEDLINE=96197133; PubMed=8626012;
RX Martin P.T., Kaufman S.J., Kramer R.H., Sanez J.R.;
RT "Synaptic integrins in developing, adult, and mutant muscle: selective
association of alpha1, alpha7a, and alpha7b integrins with the
neuromuscular junction.";
RL Dev. Biol. 174:125-139(1996).
-1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
SITES OF SECONDARY FIBER FORMATION. IT IS INVOLVED IN THE
MAINTENANCE OF THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR
ANCHORAGE, VIABILITY AND FUNCTIONAL INTEGRITY. ISOFORMS ALPHA-7X2B
AND ALPHA-7X1B PROMOTE MYOBLAST MIGRATION ON LAMININ 1 AND LAMININ
2/4, BUT ALPHA-7X1B IS LESS ACTIVE ON LAMININ 1 (IN VITRO).
-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
-1- ALTERNATIVE PRODUCTS: AT LEAST 12 ISOFORMS: ALPHA-7X1A, ALPHA-
7X1B, ALPHA-7X1DA, ALPHA-7X1DB, ALPHA-7X2A, ALPHA-7X2B, ALPHA-
7X2DA, ALPHA-7X2DB, ALPHA-7X1X2A, ALPHA-7X1X2B (SHOWN HERE),
ALPHA-7X1X2DA AND ALPHA-7X1X2DB; MAY BE PRODUCED BY ALTERNATIVE
SPLICING. THERE IS A COMBINATION OF AT LEAST FIVE ALTERNATIVE
SPLICED DOMAINS, THREE EXTRACELLULAR (X1, X2 AND D) AND TWO
CYTOPLASMIC (A AND B). A THIRD POTENTIAL ALTERNATIVE SPLICED
GENERATED ISOFORM X2C SHOWS FUNCTION. SO FAR DETECTED ARE ISOFORMS
ALPHA-7X1A, ALPHA-7X2B (MAJOR), ALPHA-7X2DB (MINOR) AND MINOR
ISOFORMS CONTAINING SEGMENT X1X2.
-1- TISSUE SPECIFICITY: ISOFORMS CONTAINING SEGMENT A ARE
PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLE. ISOFORMS CONTAINING
SEGMENT B ARE ABUNDANTLY EXPRESSED IN SKELETAL MUSCLE, MODERATELY
IN CARDIAC MUSCLE, SMALL INTESTINE, COLON, OVARY AND PROSTATE AND
WEAKLY IN LUNG AND TESTES. ISOFORMS CONTAINING SEGMENT X2D ARE

CC EXPRESSED AT LOW LEVELS IN FETAL AND ADULT SKELETAL MUSCLE AND IN
CC CARDIAC MUSCLE, BUT ARE NOT DETECTED IN MYOBLASTS AND MYOTUBES. IN
CC MUSCLE FIBERS ISOFORMS CONTAINING SEGMENT A AND B ARE EXPRESSED AT
CC MYOTENDONIC AND NEUROMUSCULAR JUNCTIONS; ISOFORMS CONTAINING
CC SEGMENT C ARE EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND AT
CC EXTRASYNAPTIC SITES. ISOFORMS CONTAINING SEGMENTS X1 OR X2 OR, AT
CC LOW LEVELS, X1X2 ARE EXPRESSED IN FETAL AND ADULT SKELETAL MUSCLE
CC (MYOBLASTS AND MYOTUBES) AND CARDIAC MUSCLE.
CC DEVELOPMENTAL STAGE: IN RENEMING INTESTINAL EPITHELIUM, EXPRESSION
CC OF ISOFORMS CONTAINING SEGMENT B CORRELATES WITH THE ONSET OF
CC ENTEROCYTIC DIFFERENTIATION.
CC -1- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES OF THE EXTRACELLULAR
CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).
CC -1- DISEASE: DEFECTS IN ITGA7 ARE ASSOCIATED WITH A FORM OF CONGENITAL
CC MYOPATHY: A GROUP OF HETEROGENEOUS MUSCLE DISORDERS WHICH ARE
CC THOUGHT TO RESULT FROM ABNORMAL MUSCLE DEVELOPMENT. MUSCLE
CC WEAKNESS IS EITHER NON-PROGRESSIVE OR SLOWLY PROGRESSIVE AND
CC APPARENT FROM BIRTH OR EARLY INFANCY.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; AF032108; AAC39708.1; -
CC DR EMBL; AF052050; AAC18968.1; -
CC DR EMBL; AF072133; AAC80458.1; -
CC DR EMBL; AJ228836; CAB41534.1; -
CC DR EMBL; AJ228837; CAB41534.1; JOINED.
CC DR EMBL; AJ228838; CAB41534.1; JOINED.
CC DR EMBL; AJ228839; CAB41534.1; JOINED.
CC DR EMBL; AJ228840; CAB41534.1; JOINED.
CC DR EMBL; AJ228842; CAB41534.1; JOINED.
CC DR EMBL; AJ228843; CAB41534.1; JOINED.
CC DR EMBL; AJ228844; CAB41534.1; JOINED.
CC DR EMBL; AJ228845; CAB41534.1; JOINED.
CC DR EMBL; AJ228846; CAB41534.1; JOINED.
CC DR EMBL; AJ228847; CAB41534.1; JOINED.
CC DR EMBL; AJ228848; CAB41534.1; JOINED.
CC DR EMBL; AJ228850; CAB41534.1; JOINED.
CC DR EMBL; AJ228851; CAB41534.1; JOINED.
CC DR EMBL; AJ228852; CAB41534.1; JOINED.
CC DR EMBL; AJ228853; CAB41534.1; JOINED.
CC DR EMBL; AJ228854; CAB41534.1; JOINED.
CC DR EMBL; AJ228855; CAB41534.1; JOINED.
CC DR EMBL; AJ228857; CAB41534.1; JOINED.
CC DR EMBL; AJ228858; CAB41534.1; JOINED.
CC DR EMBL; AJ228859; CAB41534.1; JOINED.
CC DR EMBL; AJ228860; CAB41534.1; JOINED.
CC DR EMBL; AJ228862; CAB41534.1; JOINED.
CC DR EMBL; AJ228863; CAB41535.1; -
CC DR EMBL; AJ228867; CAB41535.1; JOINED.
CC DR EMBL; AJ228868; CAB41535.1; JOINED.
CC DR EMBL; AJ228869; CAB41535.1; JOINED.
CC DR EMBL; AJ228870; CAB41535.1; JOINED.
CC DR EMBL; AJ228871; CAB41535.1; JOINED.
CC DR EMBL; AJ228872; CAB41535.1; JOINED.
CC DR EMBL; AJ228873; CAB41535.1; JOINED.
CC DR EMBL; AJ228874; CAB41535.1; JOINED.
CC DR EMBL; AJ228875; CAB41535.1; JOINED.
CC DR EMBL; AJ228876; CAB41535.1; JOINED.
CC DR EMBL; AJ228877; CAB41535.1; JOINED.
CC DR EMBL; AJ228878; CAB41535.1; JOINED.
CC DR EMBL; AJ228879; CAB41535.1; JOINED.
CC DR EMBL; AJ228880; CAB41535.1; JOINED.
CC DR EMBL; AJ228881; CAB41535.1; JOINED.
CC DR EMBL; AJ228882; CAB41535.1; JOINED.
CC DR EMBL; AJ228883; CAB41535.1; JOINED.
CC DR EMBL; AJ228884; CAB41535.1; JOINED.
CC DR EMBL; AJ228885; CAB41535.1; JOINED.
CC DR EMBL; AJ228886; CAB41535.1; JOINED.
CC DR EMBL; AJ228887; CAB41535.1; JOINED.
CC DR EMBL; AJ228888; CAB41535.1; JOINED.
CC DR EMBL; AJ228889; CAB41535.1; JOINED.
CC DR EMBL; AJ228890; CAB41535.1; JOINED.
CC DR EMBL; AJ228891; CAB41535.1; JOINED.
CC DR EMBL; AJ228892; CAB41535.1; JOINED.
CC DR EMBL; AJ228893; CAB41535.1; JOINED.
CC DR EMBL; AJ228894; CAB41535.1; JOINED.
CC DR EMBL; AJ228895; CAB41535.1; JOINED.
CC DR EMBL; AJ228896; CAB41535.1; JOINED.
CC DR EMBL; AJ228897; CAB41535.1; JOINED.
CC DR EMBL; AJ228898; CAB41535.1; JOINED.
CC DR EMBL; AJ228899; CAB41535.1; JOINED.
CC DR EMBL; AJ228900; CAB41535.1; JOINED.
CC DR EMBL; AJ228901; CAB41535.1; JOINED.
CC DR EMBL; AJ228902; CAB41535.1; JOINED.
CC DR EMBL; AJ228903; CAB41535.1; JOINED.
CC DR EMBL; AJ228904; CAB41535.1; JOINED.
CC DR EMBL; AJ228905; CAB41535.1; JOINED.
CC DR EMBL; AJ228906; CAB41535.1; JOINED.
CC DR EMBL; AJ228907; CAB41535.1; JOINED.
CC DR EMBL; AJ228908; CAB41535.1; JOINED.
CC DR EMBL; AJ228909; CAB41535.1; JOINED.
CC DR EMBL; AJ228910; CAB41535.1; JOINED.
CC DR EMBL; AJ228911; CAB41535.1; JOINED.
CC DR EMBL; AJ228912; CAB41535.1; JOINED.
CC DR EMBL; AJ228913; CAB41535.1; JOINED.
CC DR EMBL; AJ228914; CAB41535.1; JOINED.
CC DR EMBL; AJ228915; CAB41535.1; JOINED.
CC DR EMBL; AJ228916; CAB41535.1; JOINED.
CC DR EMBL; AJ228917; CAB41535.1; JOINED.
CC DR EMBL; AJ228918; CAB41535.1; JOINED.
CC DR EMBL; AJ228919; CAB41535.1; JOINED.
CC DR EMBL; AJ228920; CAB41535.1; JOINED.
CC DR EMBL; AJ228921; CAB41535.1; JOINED.
CC DR EMBL; AJ228922; CAB41535.1; JOINED.
CC DR EMBL; AJ228923; CAB41535.1; JOINED.
CC DR EMBL; AJ228924; CAB41535.1; JOINED.
CC DR EMBL; AJ228925; CAB41535.1; JOINED.
CC DR EMBL; AJ228926; CAB41535.1; JOINED.
CC DR EMBL; AJ228927; CAB41535.1; JOINED.
CC DR EMBL; AJ228928; CAB41535.1; JOINED.
CC DR EMBL; AJ228929; CAB41535.1; JOINED.
CC DR EMBL; AJ228930; CAB41535.1; JOINED.
CC DR EMBL; AJ228931; CAB41535.1; JOINED.
CC DR EMBL; AJ228932; CAB41535.1; JOINED.
CC DR EMBL; AJ228933; CAB41535.1; JOINED.
CC DR EMBL; AJ228934; CAB41535.1; JOINED.
CC DR EMBL; AJ228935; CAB41535.1; JOINED.
CC DR EMBL; AJ228936; CAB41535.1; JOINED.
CC DR EMBL; AJ228937; CAB41535.1; JOINED.
CC DR EMBL; AJ228938; CAB41535.1; JOINED.
CC DR EMBL; AJ228939; CAB41535.1; JOINED.
CC DR EMBL; AJ228940; CAB41535.1; JOINED.
CC DR EMBL; AJ228941; CAB41535.1; JOINED.
CC DR EMBL; AJ228942; CAB41535.1; JOINED.
CC DR EMBL; AJ228943; CAB41535.1; JOINED.
CC DR EMBL; AJ228944; CAB41535.1; JOINED.
CC DR EMBL; AJ228945; CAB41535.1; JOINED.
CC DR EMBL; AJ228946; CAB41535.1; JOINED.
CC DR EMBL; AJ228947; CAB41535.1; JOINED.
CC DR EMBL; AJ228948; CAB41535.1; JOINED.
CC DR EMBL; AJ228949; CAB41535.1; JOINED.
CC DR EMBL; AJ228950; CAB41535.1; JOINED.
CC DR EMBL; AJ228951; CAB41535.1; JOINED.
CC DR EMBL; AJ228952; CAB41535.1; JOINED.
CC DR EMBL; AJ228953; CAB41535.1; JOINED.
CC DR EMBL; AJ228954; CAB41535.1; JOINED.
CC DR EMBL; AJ228955; CAB41535.1; JOINED.
CC DR EMBL; AJ228956; CAB41535.1; JOINED.
CC DR EMBL; AJ228957; CAB41535.1; JOINED.
CC DR EMBL; AJ228958; CAB41535.1; JOINED.
CC DR EMBL; AJ228959; CAB41535.1; JOINED.
CC DR EMBL; AJ228960; CAB41535.1; JOINED.
CC DR EMBL; AJ228961; CAB41535.1; JOINED.
CC DR EMBL; AJ228962; CAB41535.1; JOINED.
CC DR EMBL; AJ228963; CAB41535.1; JOINED.
CC DR EMBL; AJ228964; CAB41535.1; JOINED.
CC DR EMBL; AJ228965; CAB41535.1; JOINED.
CC DR EMBL; AJ228966; CAB41535.1; JOINED.
CC DR EMBL; AJ228967; CAB41535.1; JOINED.
CC DR EMBL; AJ228968; CAB41535.1; JOINED.
CC DR EMBL; AJ228969; CAB41535.1; JOINED.
CC DR EMBL; AJ228970; CAB41535.1; JOINED.
CC DR EMBL; AJ228971; CAB41535.1; JOINED.
CC DR EMBL; AJ228972; CAB41535.1; JOINED.
CC DR EMBL; AJ228973; CAB41535.1; JOINED.
CC DR EMBL; AJ228974; CAB41535.1; JOINED.
CC DR EMBL; AJ228975; CAB41535.1; JOINED.
CC DR EMBL; AJ228976; CAB41535.1; JOINED.
CC DR EMBL; AJ228977; CAB41535.1; JOINED.
CC DR EMBL; AJ228978; CAB41535.1; JOINED.
CC DR EMBL; AJ228979; CAB41535.1; JOINED.
CC DR EMBL; AJ228980; CAB41535.1; JOINED.
CC DR EMBL; AJ228981; CAB41535.1; JOINED.
CC DR EMBL; AJ228982; CAB41535.1; JOINED.
CC DR EMBL; AJ228983; CAB41535.1; JOINED.
CC DR EMBL; AJ228984; CAB41535.1; JOINED.
CC DR EMBL; AJ228985; CAB41535.1; JOINED.
CC DR EMBL; AJ228986; CAB41535.1; JOINED.
CC DR EMBL; AJ228987; CAB41535.1; JOINED.
CC DR EMBL; AJ228988; CAB41535.1; JOINED.
CC DR EMBL; AJ228989; CAB41535.1; JOINED.
CC DR EMBL; AJ228990; CAB41535.1; JOINED.
CC DR EMBL; AJ228991; CAB41535.1; JOINED.
CC DR EMBL; AJ228992; CAB41535.1; JOINED.
CC DR EMBL; AJ228993; CAB41535.1; JOINED.
CC DR EMBL; AJ228994; CAB41535.1; JOINED.
CC DR EMBL; AJ228995; CAB41535.1; JOINED.
CC DR EMBL; AJ228996; CAB41535.1; JOINED.
CC DR EMBL; AJ228997; CAB41535.1; JOINED.
CC DR EMBL; AJ228998; CAB41535.1; JOINED.
CC DR EMBL; AJ228999; CAB41535.1; JOINED.
CC DR EMBL; AJ229000; CAB41535.1; JOINED.
CC DR EMBL; AJ229001; CAB41535.1; JOINED.
CC DR EMBL; AJ229002; CAB41535.1; JOINED.
CC DR EMBL; AJ229003; CAB41535.1; JOINED.
CC DR EMBL; AJ229004; CAB41535.1; JOINED.
CC DR EMBL; AJ229005; CAB41535.1; JOINED.
CC DR EMBL; AJ229006; CAB41535.1; JOINED.
CC DR EMBL; AJ229007; CAB41535.1; JOINED.
CC DR EMBL; AJ229008; CAB41535.1; JOINED.
CC DR EMBL; AJ229009; CAB41535.1; JOINED.
CC DR EMBL; AJ229010; CAB41535.1; JOINED.
CC DR EMBL; AJ229011; CAB41535.1; JOINED.
CC DR EMBL; AJ229012; CAB41535.1; JOINED.
CC DR EMBL; AJ229013; CAB41535.1; JOINED.
CC DR EMBL; AJ229014; CAB41535.1; JOINED.
CC DR EMBL; AJ229015; CAB41535.1; JOINED.
CC DR EMBL; AJ229016; CAB41535.1; JOINED.
CC DR EMBL; AJ229017; CAB41535.1; JOINED.
CC DR EMBL; AJ229018; CAB41535.1; JOINED.
CC DR EMBL; AJ229019; CAB41535.1; JOINED.
CC DR EMBL; AJ229020; CAB41535.1; JOINED.
CC DR EMBL; AJ229021; CAB41535.1; JOINED.
CC DR EMBL; AJ229022; CAB41535.1; JOINED.
CC DR EMBL; AJ229023; CAB41535.1; JOINED.
CC DR EMBL; AJ229024; CAB41535.1; JOINED.
CC DR EMBL; AJ229025; CAB41535.1; JOINED.
CC DR EMBL; AJ229026; CAB41535.1; JOINED.
CC DR EMBL; AJ229027; CAB41535.1; JOINED.
CC DR EMBL; AJ229028; CAB41535.1; JOINED.
CC DR EMBL; AJ229029; CAB41535.1; JOINED.
CC DR EMBL; AJ229030; CAB41535.1; JOINED.
CC DR EMBL; AJ229031; CAB41535.1; JOINED.
CC DR EMBL; AJ229032; CAB41535.1; JOINED.
CC DR EMBL; AJ229033; CAB41535.1; JOINED.
CC DR EMBL; AJ229034; CAB41535.1; JOINED.
CC DR EMBL; AJ229035; CAB41535.1; JOINED.
CC DR EMBL; AJ229036; CAB41535.1; JOINED.
CC DR EMBL; AJ229037; CAB41535.1; JOINED.
CC DR EMBL; AJ229038; CAB41535.1; JOINED.
CC DR EMBL; AJ229039; CAB41535.1; JOINED.
CC DR EMBL; AJ229040; CAB41535.1; JOINED.
CC DR EMBL; AJ229041; CAB41535.1; JOINED.
CC DR EMBL; AJ229042; CAB41535.1; JOINED.
CC DR EMBL; AJ229043; CAB41535.1; JOINED.
CC DR EMBL; AJ229044; CAB41535.1; JOINED.
CC DR EMBL; AJ229045; CAB41535.1; JOINED.
CC DR EMBL; AJ229046; CAB41535.1; JOINED.
CC DR EMBL; AJ229047; CAB41535.1; JOINED.
CC DR EMBL; AJ229048; CAB41535.1; JOINED.
CC DR EMBL; AJ229049; CAB41535.1; JOINED.
CC DR EMBL; AJ229050; CAB41535.1; JOINED.
CC DR EMBL; AJ229051; CAB41535.1; JOINED.
CC DR EMBL; AJ229052; CAB41535.1; JOINED.
CC DR EMBL; AJ229053; CAB41535.1; JOINED.
CC DR EMBL; AJ229054; CAB41535.1; JOINED.
CC DR EMBL; AJ229055; CAB41535.1; JOINED.
CC DR EMBL; AJ229056; CAB41535.1; JOINED.
CC DR EMBL; AJ229057; CAB41535.1; JOINED.
CC DR EMBL; AJ229058; CAB41535.1; JOINED.
CC DR EMBL; AJ229059; CAB41535.1; JOINED.
CC DR EMBL; AJ229060; CAB41535.1; JOINED.
CC DR EMBL; AJ229061; CAB41535.1; JOINED.
CC DR EMBL; AJ229062; CAB41535.1; JOINED.
CC DR EMBL; AJ229063; CAB41535.1; JOINED.
CC DR EMBL; AJ229064; CAB41535.1; JOINED.
CC DR EMBL; AJ229065; CAB41535.1; JOINED.
CC DR EMBL; AJ229066; CAB41535.1; JOINED.
CC DR EMBL; AJ229067; CAB41535.1; JOINED.
CC DR EMBL; AJ229068; CAB41535.1; JOINED.
CC DR EMBL; AJ229069; CAB41535.1; JOINED.
CC DR EMBL; AJ229070; CAB41535.1; JOINED.
CC DR EMBL; AJ229071; CAB41535.1; JOINED.
CC DR EMBL; AJ229072; CAB41535.1; JOINED.
CC DR EMBL; AJ229073; CAB41535.1; JOINED.
CC DR EMBL; AJ229074; CAB41535.1; JOINED.
CC DR EMBL; AJ229075; CAB41535.1; JOINED.
CC DR EMBL; AJ229076; CAB41535.1; JOINED.
CC DR EMBL; AJ229077; CAB41535.1; JOINED.
CC DR EMBL; AJ229078; CAB41535.1; JOINED.
CC DR EMBL; AJ229079; CAB41535.1; JOINED.
CC DR EMBL; AJ229080; CAB41535.1; JOINED.
CC DR EMBL; AJ229081; CAB41535.1; JOINED.
CC DR EMBL; AJ229082; CAB41535.1; JOINED.
CC DR EMBL; AJ229083; CAB41535.1; JOINED.
CC DR EMBL; AJ229084; CAB41535.1; JOINED.
CC DR EMBL; AJ229085; CAB41535.1; JOINED.
CC DR EMBL; AJ229086; CAB41535.1; JOINED.
CC DR EMBL; AJ229087; CAB41535.1; JOINED.
CC DR EMBL; AJ229088; CAB41535.1; JOINED.
CC DR EMBL; AJ229089; CAB41535.1; JOINED.
CC DR EMBL; AJ229090; CAB41535.1; JOINED.
CC DR EMBL; AJ229091; CAB41535.1; JOINED.
CC DR EMBL; AJ229092; CAB41535.1; JOINED.
CC DR EMBL; AJ229093; CAB41535.1; JOINED.
CC DR EMBL; AJ229094; CAB41535.1; JOINED.
CC DR EMBL; AJ229095; CAB41535.1; JOINED.
CC DR EMBL; AJ229096; CAB41535.1; JOINED.
CC DR EMBL; AJ229097; CAB41535.1; JOINED.
CC DR EMBL; AJ229098; CAB41535.1; JOINED.
CC DR EMBL; AJ229099; CAB41535.1; JOINED.
CC DR EMBL; AJ229100; CAB41535.1; JOINED.
CC DR EMBL; AJ229101; CAB41535.1; JOINED.
CC DR EMBL; AJ229102; CAB41535.1; JOINED.
CC DR EMBL; AJ229103; CAB41535.1; JOINED.
CC DR EMBL; AJ229104; CAB41535.1; JOINED.
CC DR EMBL; AJ229105; CAB41535.1; JOINED.
CC DR EMBL; AJ229106; CAB41535.1; JOINED.
CC DR EMBL; AJ229107; CAB41535.1; JOINED.
CC DR EMBL; AJ229108; CAB41535.1; JOINED.
CC DR EMBL; AJ229109; CAB41535.1; JOINED.
CC DR EMBL; AJ229110; CAB41535.1; JOINED.
CC DR EMBL; AJ229111; CAB41535.1; JOINED.
CC DR EMBL; AJ229112; CAB41535.1; JOINED.
CC DR EMBL; AJ229113; CAB41535.1; JOINED.
CC DR EMBL; AJ229114; CAB41535.1; JOINED.
CC DR EMBL; AJ229115; CAB41535.1; JOINED.
CC DR EMBL; AJ229116; CAB41535.1; JOINED.
CC DR EMBL; AJ229117; CAB41535.1; JOINED.
CC DR EMBL; AJ229118; CAB41535.1; JOINED.
CC DR EMBL; AJ229119; CAB41535.1; JOINED.
CC DR EMBL; AJ229120; CAB41535.1; JOINED.
CC DR EMBL; AJ229121; CAB41535.1; JOINED.
CC DR EMBL; AJ229122; CAB41535.1; JOINED.
CC DR EMBL; AJ229123; CAB41535.1; JOINED.
CC DR EMBL; AJ229124; CAB41535.1; JOINED.
CC DR EMBL; AJ229125; CAB41535.1; JOINED.
CC DR EMBL; AJ229126; CAB41535.1; JOINED.
CC DR EMBL; AJ229127; CAB41535.1; JOINED.
CC DR EMBL; AJ229128; CAB41535.1; JOINED.
CC DR EMBL; AJ229129; CAB41535.1; JOINED.
CC DR EMBL; AJ229130; CAB41535.1; JOINED.
CC DR EMBL; AJ229131; CAB41535.1; JOINED.
CC DR EMBL; AJ229132; CAB41535.1; JOINED.
CC DR EMBL; AJ229133; CAB41535.1; JOINED.
CC DR EMBL; AJ229134; CAB41535.1; JOINED.
CC DR EMBL; AJ229135; CAB41535.1; JOINED.
CC DR EMBL; AJ229136; CAB41535.1; JOINED.
CC DR EMBL; AJ229137; CAB41535.1; JOINED.
CC DR EMBL; AJ229138; CAB41535.1; JOINED.
CC DR EMBL; AJ229139; CAB41535.1; JOINED.
CC DR EMBL; AJ229140; CAB41535.1; JOINED.
CC DR EMBL; AJ229141; CAB41535.1; JOINED.
CC DR EMBL; AJ229142; CAB41535.1; JOINED.
CC DR EMBL; AJ229143; CAB41535.1; JOINED.
CC DR EMBL; AJ229144; CAB41535.1; JOINED.
CC DR EMBL; AJ229145; CAB41535.1; JOINED.
CC DR EMBL; AJ229146; CAB41535.1; JOINED.
CC DR EMBL; AJ229147; CAB41535.1; JOINED.
CC DR EMBL; AJ229148; CAB41535.1; JOINED.
CC DR EMBL; AJ229149; CAB41535.1; JOINED.
CC DR EMBL; AJ229150; CAB41535.1; JOINED.
CC DR EMBL; AJ229151; CAB41535.1; JOINED.
CC DR EMBL; AJ229152; CAB41535.1; JOINED.
CC DR EMBL; AJ229153; CAB41535.1; JOINED.
CC DR EMBL; AJ229154; CAB41535.1; JOINED.
CC DR EMBL; AJ229155; CAB41535.1; JOINED.
CC DR EMBL; AJ229156; CAB41535.1; JOINED.
CC DR EMBL; AJ229157; CAB41535.1; JOINED.
CC DR EMBL; AJ229158; CAB41535.1; JOINED.
CC DR EMBL; AJ229159; CAB41535.1; JOINED.
CC DR EMBL; AJ229160; CAB41535.1; JOINED.
CC DR EMBL; AJ229161; CAB41535.1; JOINED.
CC DR EMBL; AJ229162; CAB41535.1; JOINED.
CC DR EMBL; AJ229163; CAB41535.1; JOINED.
CC DR EMBL; AJ229164; CAB41535.1; JOINED.
CC DR EMBL; AJ229165; CAB41535.1; JOINED.
CC DR EMBL; AJ229166; CAB41535.1; JOINED.
CC DR EMBL; AJ229167; CAB41535.1; JOINED.
CC DR EMBL; AJ229168; CAB41535.1; JOINED.
CC DR EMBL; AJ229169; CAB41535.1; JOINED.
CC DR EMBL; AJ229170; CAB41535.1; JOINED.
CC DR EMBL; AJ229171; CAB41535.1; JOINED.
CC DR EMBL; AJ229172; CAB41535.1; JOINED.
CC DR EMBL; AJ229173; CAB41535.1; JOINED.
CC DR EMBL; AJ229174; CAB41535.1; JOINED.
CC DR EMBL; AJ229175; CAB41535.1; JOINED.
CC DR EMBL; AJ229176; CAB41535.1; JOINED.
CC DR EMBL; AJ229177; CAB41535.1; JOINED.
CC DR EMBL; AJ229178; CAB41535.1; JOINED.
CC DR EMBL; AJ229179; CAB41535.1; JOINED.
CC DR EMBL; AJ229180; CAB41535.1; JOINED.
CC DR EMBL; AJ229181; CAB41535.1; JOINED.
CC DR EMBL; AJ229182; CAB41535.1; JOINED.
CC DR EMBL; AJ229183; CAB41535.1; JOINED.
CC DR EMBL; AJ229184; CAB41535.1; JOINED.
CC DR EMBL; AJ229185; CAB41535.1; JOINED.
CC DR EMBL; AJ229186; CAB41535.1; JOINED.
CC DR EMBL; AJ229187; CAB41535.1; JOINED.
CC DR EMBL; AJ229188; CAB41535.1; JOINED.
CC DR EMBL; AJ229189; CAB41535.1; JOINED.
CC DR EMBL; AJ229190; CAB41535.1; JOINED.
CC DR EMBL; AJ229191; CAB41535.1; JOINED.
CC DR EMBL; AJ229192; CAB41535.1; JOINED.
CC DR EMBL; AJ229193; CAB41535.1; JOINED.
CC DR EMBL; AJ229194; CAB41535.1; JOINED.
CC DR EMBL; AJ229195; CAB41535.1; JOINED.
CC DR EMBL; AJ229196; CAB41535.1; JOINED.
CC DR EMBL; AJ229197; CAB41535.1; JOINED.
CC DR EMBL; AJ229198; CAB41535.1; JOINED.
CC DR EMBL; AJ229199; CAB41535.1; JOINED.
CC DR

```

OY 7 LVLALLV 14
DB 1093 LVLALLV 1100

RESULT 9
YEDD_SALTY STANDARD; PRT; 141 AA.
ID YEDD_SALTY
AC 006399;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN YEDD.
GN YEDD.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmonella.
ON NCBI_TaxID=602;
RX 111
RP SEQUENCE FROM N.A.
RC STRAIN=SV1103;
RA MEDLINE=93381452; PubMed=8371104;
RA Raha M., Kihara M., Kawagishi I., Macnab R.M.;
RA Organization of the Escherichia coli and Salmonella typhimurium
RA chromosomes between flagellar regions IIA and IIB, including a
RA large non-coding region.
RT J. Gen. Microbiol. 139:1401-1407(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13280; AAA1971.1;
DR StGene: SC10420; yedd.
KW Hypothetical protein.
SQ SEQUENCE 141 AA; 15465 MW; A37E9E11ADC8BD9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 141;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 GALVIA 11
DB 8 GALVIA 14

RESULT 10
LYC2_PIG STANDARD; PRT; 146 AA.
ID LYC2_PIG
AC P12068;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYSOZYME C-2 PRECURSOR (EC 3.2.1.17) (1,4-BETA-N-ACETYLMURAMIDASE C).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
ON NCBI_TaxID=9823;
RX 111
RP SEQUENCE FROM N.A.
RC MEDLINE=96292525; PubMed=8728388;
RA Yu M., Irwin D.M.;
RA "Evolution of stomach lysozyme: the pig lysozyme gene.";
RA Mol. Phylogenet. Evol. 5:298-308(1996).
RN [2]
RP SEQUENCE OF 19-146.
RC TISSUE=Stomach;

```

```

RX MEDLINE=89362490; PubMed=2504928;
RA Jolles J., Jolles P., Bowman B.H., Prager E.M., Stewart C.-B.,
RA Wilson A.C.;
RT "Epistatic evolution in the stomach lysozymes of ruminants.";
RT J. Mol. Evol. 28:528-535(1989).
CC -----
CC -1- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION: THOSE
CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-
CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOGENE-
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES BETWEEN
CC N-ACETYL-D-GLUCOSAMINE AND N-ACETYLMURAMIC ACID IN PEPTIDOGLYCAN
CC HETEROPOLYMERS OF THE PROKARYOTES CELL WALLS.
CC -1- SUBUNIT: MONOMER.
CC -1- MISCELLANEOUS: LYSOZYMES C ARE CAPABLE OF BOTH HYDROLYSIS &
CC TRANSGLYCOSYLATION: THEY SHOW ALSO A SLIGHT ESTERASE ACTIVITY.
CC THEY ACT RAPIDLY ON BOTH PEPTIDE-SUBSTITUTED & UNSUBSTITUTED
CC PEPTIDOGLYCAN & SLOWLY, ON CHITIN OLIGOSACCHARIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28757; AAB16862.1;
DR HSSP: P00695; 2LHM.
DR InterPro: IPR001916; Lactalbum_Lysozyme.
DR Pfam: PF00062; lys; 1.
DR PRINTS: PRO0135; LYZLACT.
DR PRINTS: PRO0137; LYSOZYME.
DR SMART: SM00263; LYZ1; 1.
DR PROSITE: PS00128; LACTALBUMIN_LYSOZYME; 1.
KW Hydrolyase; Glycosidase; Bacteriolytic enzyme; Digestion; Stomach;
KW Signal.
FT CHAIN 1 18
FT SIGNAL 1 18
FT DISULEID 24 146
FT DISULEID 24 144
FT DISULEID 48 132
FT DISULEID 81 97
FT DISULEID 93 111
FT ACT_SITE 53 53
FT ACT_SITE 69 69
FT ACT_SITE 69 69
SQ SEQUENCE 146 AA; 16484 MW; C2986F5CAAF49F4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 146;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LVLALLV 13
DB 4 LVLALLV 10

RESULT 11
LYC3_PIG STANDARD; PRT; 148 AA.
ID LYC3_PIG
AC P12069;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYSOZYME C-3 PRECURSOR (EC 3.2.1.17) (1,4-BETA-N-ACETYLMURAMIDASE C).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
ON NCBI_TaxID=9823;
RX 111
RP SEQUENCE FROM N.A.
RA Echeverdu Z.O., Nevills M., Bixby J.A., Roberts R.M., Trout W.E.;
RA Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP SEQUENCE OF 19-148.
RC TISSUE=Stomach;
RX MEDLINE=69362490; PubMed=2504928;
RA Jolles J., Jolles P., Bowman B.H., Prager E.M., Stewart C.B.,
RA Wilson A.C.;
RT "Episodic evolution in the stomach Lysozymes of ruminants.";
RJ Mol. Evol. 28:528-535(1989).
CC -1- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION; THOSE
CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-
CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOAGENTS.
CC -1- CATALYTIC ACTIVITY: HYDROLASES OF THE 1,4-BETA-LINKAGES BETWEEN
CC N-ACETYL-D-GLUCOSAMINE AND N-ACETYL-MURAMIC ACID IN PEPTIDOGLYCAN
CC HETEROPOLYMERS OF THE PROKARYOTES CELL WALLS.
CC -1- SUBUNIT: MONOMER.
CC -1- MISCELLANEOUS: LYSOZYMES C ARE CAPABLE OF BOTH HYDROLYSIS &
CC TRANSGLYCOSYLATION; THEY SHOW ALSO A SLIGHT ESTERASE ACTIVITY.
CC THEY ACT RAPIDLY ON BOTH PEPTIDE-SUBSTITUTED & UNSUBSTITUTED
CC PEPTIDOGLYCAN &, SLOWLY, ON CHITIN OLIGOSACCHARIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U44435; AAA86644.1; -.
CC
CC PIR: S10047; S10047.
CC
CC HSSP; P00695; 2LHM.
CC InterPro: IPR001916; Lactalbumin_Lysozyme.
CC Pfam; PF00062; Lys; 1.
CC PRINTS; PR00135; LYLIACT.
CC PRINTS; PR00137; LYSOZYME.
CC SMART; SM00263; LY21; 1.
CC
CC PROSITE; PS00126; LACTALBUMIN_LYSOZYME; 1.
CC
CC Hydrolase; Glycosidase; Bacteriolytic enzyme; Digestion; Stomach;
CC Signal.
CC
CC FT CHAIN 1 18 LYSOZYME C-3.
CC FT DISULFID 19 148 BY SIMILARITY.
CC FT DISULFID 24 146 BY SIMILARITY.
CC FT DISULFID 48 134 BY SIMILARITY.
CC FT DISULFID 83 99 BY SIMILARITY.
CC FT DISULFID 95 113 BY SIMILARITY.
CC FT ACT_SITE 53 53 BY SIMILARITY.
CC FT ACT_SITE 71 71 BY SIMILARITY.
CC
CC SO SEQUENCE 148 AA; 16711 MW; 76AA67BFA8BD64E6 CRC64;

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE-99096913; PubMed-9878396;
RA Ihara K., Umemura T., Katagiri I., Kitajima-Ihara T., Sugiyama Y.,
RA Kimura Y., Mikhonata Y.,
RT "Evolution of the archaeal rhodopsins: evolution rate changes by gene
RT duplication and functional differentiation.";
RL J. Mol. Biol. 285:163-174(1999).
CC -1- FUNCTION: BACTERIORHODOPSIN IS A LIGHT-DRIVEN PROTON PUMP.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL OPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; AB009620; BAA75200.1; -.
CC HSSP; P02945; 1F50.
CC InterPro; IPR001425; Bac_rhodopsin.
CC Pfam; PF01036; Bac_rhodopsin; 1.
CC PRINTS; PR00251; BACTERIAL_OPSIN.
CC PROSITE; PS00950; BACTERIAL_OPSIN_1; 1.
CC PROSITE; PS00327; BACTERIAL_OPSIN_RET; 1.
CC Ion transport; Photoreceptor; Transmembrane; Retinal protein;
CC Hydrogen ion transport.
CC DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 17 37 HELIX A (POTENTIAL).
CC FT TRANSMEM 52 72 HELIX B (POTENTIAL).
CC FT TRANSMEM 91 111 HELIX C (POTENTIAL).
CC FT TRANSMEM 115 135 HELIX D (POTENTIAL).
CC FT TRANSMEM 143 163 HELIX E (POTENTIAL).
CC FT TRANSMEM 185 205 HELIX F (POTENTIAL).
CC FT TRANSMEM 213 233 HELIX G (POTENTIAL).
CC FT DOMAIN 234 250 CYTOPLASMIC (POTENTIAL).
CC FT BINDING 225 225 RETINAL CHROMOPHORE (BY SIMILARITY).
CC FT SEQUENCE 250 AA; 27041 MW; B72D3373506FD275 CRC64;
SQ
Query Match 6.3%; Score 7; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 42 LLLDLSL 48
|||||||
Db 101 LLLDLSL 107

```

RL J. Biol. Chem. 267:10225-10231(1992).
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA GLASER P., Kunst F., Anand M., Coudart M.P., Gonzales W.,
 RA Hulo M.F., Ionescu M., Lubchinsky B., Marcelino L., Moser I.,
 RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
 RA Kapoport G., Danchin A.;
 RA Bacillus subtilis genome project: cloning and sequencing of the 97
 RA kd region from 325 degrees to 333 degrees.;
 RL Mol. Microbiol. 10:371-384(1993).
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RC SEQUENCE OF 27-43, AND CHARACTERIZATION.
 RA STRAIN-23;
 RA MEDLINE-96009139; PubMed-7575098;
 RA Lemme E., Simon J., Schagger H., Kroger A.;
 RA Properties of the menaquinol oxidase (cox) and of cox deletion
 RA mutants of Bacillus subtilis.;
 RL Arch. Microbiol. 163:432-438(1995).
 CC -1- FUNCTION: CATALYZES QUINOL OXIDATION WITH THE CONCOMITANT
 CC REDUCTION OF OXYGEN TO WATER. MAJOR COMPONENT FOR ENERGY
 CC CONVERSION DURING VEGETATIVE GROWTH. SUBUNIT II TRANSFERS THE
 CC ELECTRONS FROM A QUINOL TO THE BINDING CENTER OF THE CATALYTIC
 CC SUBUNIT I.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: SOME, TO MITOCHONDRIAL OR BACTERIAL COX2 SUBUNITS.
 CC BUT LACK HEME-BINDING DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M86548; AAA22686.1; ALT_INIT.
 DR EMBL; X73124; CA51593.1; ALT_INIT.
 DR EMBL; 299123; CAB1843.1; ALT_INIT.
 DR PIR; A38129; A38129.
 DR HSSP; P18400; ICYM.
 DR Subtilist; B610583; goxa.
 DR InterPro; IPR001505; COX2.
 DR InterPro; IPR002429; Cyl_c-cox-2.
 DR Pfam; PF00116; COX2.1.
 DR PRINTS; PR01166; CYCOXIDASE1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 DR Oxidoreductase; Electron transport; Transmembrane; Signal;
 KW Lipoprotein; Complete proteome.
 FT SIGNAL 1 26
 FT CHAIN 27 322
 FT LIPID 27 27
 FT TRANSMEM 50 70
 FT TRANSMEM 92 112
 FT CONFLICT 39 39
 FT SEQUENCE 322 AA; 36353 MW; C701B14076B7A8D CRC64;
 V -> VD (IN REF. X).
 Query Match 6.3%; Score 7; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 LTVALL 13
 Db 12 LTVALL 18

RESULT 14
 ID C1B2_CAVPO STANDARD; PRT; 332 AA.
 AC 090221;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE T-CELL SURFACE GLYCOPROTEIN CD1B2 PRECURSOR (CD1-B2 ANTIGEN).
 GN CD1B2.
 OS Eukarya; porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae; Cavia.
 OX NCBI_Taxid:10141;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-HARTLEY, AND NIH 2; TISSUE=Thymus;
 RX MEDLINE-20021845; PubMed-10553074;
 RA Dascher C.C., Hirumatsu K., Naylor J.W., Brauer P.P., Brown K.A.,
 RA Storey J.R., Behar S.M., Kawasaki E.S., Porcelli S.A., Brenner M.B.,
 RA Leclaire K.P.;
 RT Conservation of a CD1 multigene family in the guinea pig.;
 RL J. Immunol. 163:5478-5488(1999).
 CC -1- FUNCTION: NOT KNOWN.
 CC -1- SUBUNIT: ASSOCIATES NON-COVALENTLY WITH BETA-2-MICROGLOBULIN (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF145484; AAF12739.1;
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00407; IgC1.1.
 DR SMART; SM00407; IgC1.1.
 DR Glycoprotein; Signal; Transmembrane; Immunoglobulin domain;
 KW Multigene family
 FT SIGNAL 1 17
 FT CHAIN 18 332
 FT DOMAIN 18 301
 FT TRANSMEM 302 322
 FT DOMAIN 323 332
 FT DISULFID 120 184
 FT DISULFID 224 279
 FT CARBOHYD 38 38
 FT CARBOHYD 146 146
 FT CARBOHYD 146 146
 FT CARBOHYD 297 297
 FT SEQUENCE 332 AA; 37384 MW; 11D9360657AD3A CRC64;
 V -> VD (IN REF. X).
 Query Match 6.3%; Score 7; DB 1; Length 332;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 7 LTVALL 13
 Db 3 LTVALL 9
 RESULT 15
 ID YE71_HAEIN STANDARD; PRT; 337 AA.
 AC 057130; 005065;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2002, 12:16:24 ; Search time 14.31 Seconds

(without alignments)
590.872 Million cell updates/sec

Title: US-09-662-784-6

Perfect score: 111

Sequence: 1 DFMRGALLVLTALVTQALGV.....CMGEAVQNTVEDLKLNTLGR 111

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	98.2	109	2	C56413
2	81	73.0	107	2	JC1127
3	8	7.2	108	2	G40149
4	8	7.2	129	2	I61187
5	8	7.2	139	1	C45051
6	8	7.2	148	2	S40148
7	8	7.2	299	2	T06960
8	8	7.2	378	2	JC5649
9	8	7.2	378	2	E82254
10	8	7.2	401	2	E82951
11	8	7.2	1106	2	S38783
12	8	7.2	1135	2	I61186
13	8	7.2	1137	2	JC5950
14	7	6.3	20	2	B53283
15	7	6.3	73	2	T20911
16	7	6.3	77	2	D83680
17	7	6.3	87	2	G46449
18	7	6.3	103	2	G75257
19	7	6.3	198	2	G59094
20	7	6.3	203	2	B75497
21	7	6.3	244	2	D75505
22	7	6.3	250	2	T50677
23	7	6.3	253	2	T46454
24	7	6.3	266	2	T36287
25	7	6.3	268	2	A84746
26	7	6.3	279	2	S75087
27	7	6.3	303	2	T00479
28	7	6.3	321	1	E69687
29	7	6.3	322	2	B83579

30	7	6.3	337	1	G64125	hemin transport pr
31	7	6.3	341	2	C72403	hypothetical prote
32	7	6.3	350	2	S75065	sensory transducti
33	7	6.3	355	2	JC1249	peroxidase (EC 1.1
34	7	6.3	364	2	S34355	peroxidase (EC 1.1
35	7	6.3	374	2	B82488	hypothetical prote
36	7	6.3	385	2	E83506	probable MFS trans
37	7	6.3	390	2	H70904	probable lprk prot
38	7	6.3	397	2	S09813	hypothetical prote
39	7	6.3	400	2	B70621	probable argp prot
40	7	6.3	400	2	T35334	probable membrane
41	7	6.3	422	2	G83503	probable MFS trans
42	7	6.3	424	2	S32488	dihydrolipeamide S
43	7	6.3	452	2	E83482	probable MFS trans
44	7	6.3	455	2	B86722	biotin carboxylase
45	7	6.3	474	2	F82412	probable aspartoki

ALIGNMENTS

RESULT 1
C56413
Major allergen Fel d1 chain 2 precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 01-Dec-2000
C:Accession: C06413; F06413
R:Morgensstern, J.P.; Griffith, I.J.; Brauer, A.W.; Rogers, B.L.; Bond, J.F.; Chapman, Proc. Natl. Acad. Sci. U.S.A. 88, 9690-9694, 1991
A:Title: Amino acid sequence of Fel d1, the major allergen of the domestic cat: prote
A:Accession: C56413
A:Reference number: A56413; M0ID:92052157
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-109 <MOR>
A:Cross-references: GB:M7341; NID:9163822; PIDN:AAC41616.1; PID:9163823
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgensstern, J.P.; Rogers, B.L. Gene 113, 263-268, 1992
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg
A:Reference number: JC1126; M0ID:92241678
A:Accession: JC1145
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 18-109 <GRI>
A:Experimental source: salivary gland
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.2% Score 109; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 7.7e-101;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MRGALLVLTALVTQALGVKAETCPFYDVFFVANGNELLDLSITKVATEPERFAMK 62
DB 1 MRGALLVLTALVTQALGVKAETCPFYDVFFVANGNELLDLSITKVATEPERFAMK 60

QY 63 KIDCVENGLISRVLDGLVMTTSSKDCMGEAVQNTVEDLKLNTLGR 111
DB 61 KIDCVENGLISRVLDGLVMTTSSKDCMGEAVQNTVEDLKLNTLGR 109

RESULT 2
JC1127
Major allergen chain 2 precursor, short form - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: JC1127
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgensstern, J.P.; Rogers, B.L. Gene 113, 263-268, 1992
A:Title: Expression and genomic structure of the genes encoding Fd1, the major allerg
A:Reference number: JC1126; M0ID:92241678

A:Accession: J01127
 A:Molecule type: DNA
 A:Residues: 1-107 <GR1>
 A:Cross-references: GB:K62478; NID:9395406; PIDN:CAA44345.1; PID:9395407
 A:Experimental source: skin
 C:Genetics:
 A:Gene: Ch2
 A:Introns: 21/1; 81/3
 C:Keywords: glycoprotein
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>
 F:50/Binding site: carbohydrate (Asn) #status predicted

Query Match 73.0%; Score 81; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 4,9e-73;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 MKGALLVALLVTOALGVKMAETCPFYDFVFAVANGNELLDLSLFVNATEPERTAMK 62
 DB 1 MKGALLVALLVTOALGVKMAETCPFYDFVFAVANGNELLDLSLFVNATEPERTAMK 60

OY 63 KIQDCVYENGILSRVLDGLVM 83
 DB 61 KIQDCVYENGILSRVLDGLVM 81

RESULT 3
 S40149
 Integrin alpha-7C chain - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 29-Sep-1999
 C:Accession: S40149
 R:Song, W.K.; Wang, W.; Sato, H.; Bjelseter, D.; Kaufman, S.
 Submitted to the EMBL Data Library July 1993
 A:Description: Expression of alpha 7 integrin cytoplasmic domains during skeletal muscle myogenesis.
 A:Reference number: S40147
 A:Accession: S40149
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-108 <SON>
 A:Cross-references: EMBL:X74294; NID:9437912; PIDN:CAA52347.1; PID:9437913
 C:Superfamily: Integrin alpha-2b chain

Query Match 7.2%; Score 8; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LTVALLV 14
 DB 79 LTVALLV 86

RESULT 4
 I61187
 Integrin - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
 C:Accession: I61187
 R:Ziober, B.L.; Vu, M.P.; Wajsb, N.; Crawford, J.; Lin, C.S.; Kramer, R.H.
 J. Biol. Chem. 268, 26773-26783, 1993
 A:Title: Alternative extracellular and cytoplasmic domains of the integrin alpha 7 subunit
 A:Reference number: A49691; MUID:94075378
 A:Accession: I61187
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-129 <RES>
 A:Cross-references: GB:L23421; NID:9431417; PIDN:AAA16598.1; PID:9450648
 C:Superfamily: Integrin alpha-2b chain

Query Match 7.2%; Score 8; DB 2; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LTVALLV 14
 DB 61 LTVALLV 68

RESULT 5
 C45051
 Lampyrin 2 precursor, long splice form - sea lamprey
 N:Alternate names: Lampyrin L-0.9-12
 N:Contains: Lampyrin 2 precursor, short splice form (Lampyrin L-0.9-10)
 C:Species: Petromyzon marinus (sea lamprey)
 C:Date: 30-Apr-1993 #sequence_revision 30-Apr-1993 #text_change 22-Jun-1999
 C:Accession: C45051; A45051
 R:Robson, P.; Wright, G.M.; Sitarz, E.; Maiti, A.; Rawat, M.; Youson, J.H.; Keeley, F.
 J. Biol. Chem. 268, 1440-1447, 1993
 A:Title: Characterization of lampyrin, an unusual matrix protein from lamprey cartilage
 A:Reference number: A45051; MUID:93123269
 A:Accession: C45051
 A:Molecule type: mRNA
 A:Residues: 1-139 <ROB>
 A:Cross-references: GB:L05925; NID:9213209; PIDN:AAA49269.1; PID:9213210
 A:Note: sequence extracted from NCBI backbone (NCBI:P122170)
 A:Accession: A45051
 A:Molecule type: mRNA
 A:Residues: 1-85,105-139 <RO2>
 A:Cross-references: GB:L05924; NID:9213207; PIDN:AAA49268.1; PID:9213208
 A:Note: sequence extracted from NCBI backbone (NCBI:P122166)
 C:Comment: These forms are encoded by a different gene than B45051; we have arbitrarily
 C:Superfamily: Lampyrin
 C:Keywords: alternative splicing; cartilage; extracellular matrix
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-139/Product: Lampyrin 2, long splice form #status predicted <MAT1>
 F:20-85,105-139/Product: Lampyrin 2, short splice form #status predicted <MAT2>

Query Match 7.2%; Score 8; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 ALTVALL 13
 DB 7 ALTVALL 14

RESULT 6
 S40148
 Integrin alpha-7A chain - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 29-Sep-1999
 C:Accession: S40148
 R:Song, W.K.; Wang, W.; Sato, H.; Bjelseter, D.; Kaufman, S.
 Submitted to the EMBL Data Library July 1993
 A:Description: Expression of alpha 7 integrin cytoplasmic domains during skeletal muscle myogenesis.
 A:Reference number: S40147
 A:Accession: S40148
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-148 <SON>
 A:Cross-references: EMBL:X74293; NID:9437910; PIDN:CAA52346.1; PID:9437911
 C:Superfamily: Integrin alpha-2b chain

Query Match 7.2%; Score 8; DB 2; Length 148;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LTVALLV 14
 DB 79 LTVALLV 86

Db 79 LVLLLV 86

RESULT 7

T06960

probable membrane protein - Cyanophora paradoxa cyanelle
N:Alternate names: hypothetical protein psc region
C:Species: cyanelle Cyanophora paradoxa
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 24-Nov-1999
C:Accession: T06960; PS0372
R:Stirewalt, V.L.; Michalowski, C.B.; Luftelhardt, W.; Bohmert, H.J.; Bryant, D.A.
submitted to the EMBL Data Library, July 1995
A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A:Reference number: Z15840
A:Accession: T06960
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-299 <STI>
A:Cross-references: EMBL:U30821; NID:g1016083; PIDN:AAA81303.1; PID:g1016216
A:Experimental source: strain Pringsheim LB555
R:Rhie, E.; Stirewalt, V.L.; Gasparich, G.E.; Bryant, D.A.
Gene 112, 123-128, 1992
A:Title: The psc genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning an
A:Reference number: J50694; MUID:92201692
A:Accession: PS0372
A:Molecule type: DNA
A:Residues: 138-299 <RHI>
A:Cross-references: GB:M86239; NID:g336634; PIDN:AAA65471.1; PID:g336637
C:Genetics:
A:Genome: cyanelle
C:Superfamily: conserved hypothetical protein HI0360
C:Keywords: cyanelle; transmembrane protein

Query Match 7.2%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 LVLLLV 15
Db 220 LVLLLV 227

RESULT 8
J5649
N-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) - Vibrio cholerae (strain non-
C:Species: Vibrio cholerae
A:Note: non-O1 strain 1148A
C>Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 05-Nov-1999
C:Accession: J5649
R:Yamano, N.; Wang, J.; Fujishima, S.
Biotech. Biotechnol. Biochem. 61, 1349-1353, 1997
A:Title: Cloning and sequencing of the genes for N-acetylglucosamine use that construct
A:Reference number: J5649; MUID:97446530
A:Accession: J5649
A:Molecule type: DNA
A:Residues: 1-378 <YAM>
A:Cross-references: DBJ:D87820; NID:g2541899; PIDN:BA22834.1; PID:d1023704; PID:g25419
A:Experimental source: strain 1148A
C:Comment: This enzyme deacetylates N-acetylglucosamine, one component of chitin, to glu
C:Genetics:
A:Gene: nagA
C:Keywords: hydrolase

Query Match 7.2%; Score 8; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 95 EAVQNTVE 102
Db 312 EAVQNTVE 319

RESULT 9

E82254

N-acetylglucosamine-6-phosphate deacetylase VC0994 [imported] - Vibrio cholerae (stra
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82254
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
charson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qu, H.; Dragol, I.; Sellers
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: E82254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <HEI>
A:Cross-references: GB:AE004181; GB:AE003852; NID:g9655454; PIDN:AAF94155.1; GSPDB:GN
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0994
A:Map position: 1

Query Match 7.2%; Score 8; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 95 EAVQNTVE 102
Db 312 EAVQNTVE 319

RESULT 10
E82951
probable MFS transporter PA5548 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E82951
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lardig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: E82951
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <STO>
A:Cross-references: GB:AE004967; GB:AE004091; NID:g9951884; PIDN:AA08933.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5548

Query Match 7.2%; Score 8; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LVLLLV 14
Db 281 LVLLLV 288

RESULT 11
S38783
integral alpha chain - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 22-Jan-1994 #sequence_revision 14-Jul-1995 #text_change 29-Sep-1999
C:Accession: S38783; S23600
R:Kauffman, S.J.
submitted to the EMBL Data Library, March 1992
A:Reference number: S38783

A:Accession: S38783
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1106 <KAU>
 A:Cross-references: EMBL:X65036; NID:g56392; PIDN:CAA6170.1; PID:g56393
 R:Song, W.K.; Wang, W.; Foster, R.F.; Bieser, D.A.; Kaufman, S.J.
 J. Cell Biol. 117, 643-657, 1992
 A:Title: H36-alpha is a novel integrin alpha chain that is developmentally regulated du
 A:Reference number: S23600; MUID:92242309
 A:Accession: S23600
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-116, 'R', 118-349, 'D', 351-804, 'R', 806, 'V', 808-1106 <SON>
 A:Cross-references: EMBL:X65036
 C:Superfamily: Integrin alpha-2b chain

Query Match 7.2%; Score 8; DB 2; Length 1106;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LVYALLV 14
 |||||

DB 1018 LVYALLV 1025

RESULT 12
 16186
 alpha-7 integrin - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
 C:Accession: 161186
 R:Zlobor, B.L.; Vu, M.P.; Waleh, N.; Crawford, J.; Lin, C.S.; Kramer, R.H.
 J. Biol. Chem. 268, 26773-26783, 1993
 A:Title: Alternative extracellular and cytoplasmic domains of the integrin alpha 7 subu
 A:Reference number: A49691; MUID:94075378
 A:Accession: 161186
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1135 <RES>
 A:Cross-references: GB:L23423; NID:g431415; PIDN:AAA16600.1; PID:g431416
 C:Superfamily: Integrin alpha-2b chain

Query Match 7.2%; Score 8; DB 2; Length 1135;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LVYALLV 14
 |||||

DB 1047 LVYALLV 1054

RESULT 13
 JC5950
 Integrin alpha-7 chain precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 31-Dec-2000
 C:Accession: JC5950; A56839; S40147
 R:Leung, E.; Lin, S.P.; Berg, R.; Yang, Y.; Ni, J.; Wang, S.; Kristensen, G.W.
 Biochem. Biophys. Res. Commun. 243, 317-325, 1998
 A:Title: A novel extracellular domain variant of the human integrin alpha 7 subunit gene
 A:Reference number: JC5950; MUID:98139911
 A:Accession: JC5950
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1137 <LEU1>
 A:Cross-references: GB:AF032108; NID:g2897115; PIDN:AAC39708.1; PID:g2897116
 A:Accession: JC5951
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-657, 733-1137 <LEU2>
 A:Cross-references: GB:AF032108

R:Wang, W.; Wu, W.; Desai, T.; Ward, D.C.; Kaufman, S.J.
 Genomics 26, 563-570, 1995
 A:Title: Localization of the alpha7 integrin gene (ITGA7) on human chromosome 12q13:
 A:Reference number: A56839
 A:Accession: A56839
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 664-666, 668-688, 690-825, 'X', 827-839, 'D', 841-946, 'H', 948, 'T', 951, 'P', 953-1
 A:Cross-references: EMBL:X74295; NID:g437781
 R:Song, W.K.; Wang, W.; Sato, H.; Bieser, D.; Kaufman, S.
 submitted to the EMBL Data Library, July 1993
 A:Description: Expression of alpha 7 integrin cytoplasmic domains during skeletal mus
 hosphases.
 A:Reference number: S40147
 A:Accession: S40147
 A:Molecule type: mRNA
 A:Residues: 1061-1137 <SON>
 A:Cross-references: EMBL:X74295; NID:g437781; PIDN:CAA52348.1; PID:g437782
 C:Genes: ITGA7
 A:Gene: GDB:ITGA7
 A:Cross-references: GDB:131465; OMIM:600536
 A:Map position: 12q13-12q13
 C:Superfamily: Integrin alpha-2b chain
 C:Keywords: glycoprotein; skeletal muscle; transmembrane protein
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:34-1137/Product: integrin alpha-7 chain, long splice form #status predicted <MNT>
 F:34-657,733-1137/Product: integrin alpha-7 chain, short splice form #status predict
 F:1038-1056/Domain: transmembrane #status predicted <TRM>
 F:86,742,945,981,1001/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.2%; Score 8; DB 2; Length 1137;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LVYALLV 14
 |||||

DB 1049 LVYALLV 1056

RESULT 14
 B53283
 major cat allergen Fel d I beta chain - cat (fragment)
 C:Species: Felis silvestris catus (domestic cat)
 C:Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 12-May-1994
 C:Accession: B53283
 R:Bufford, O.A.; Carreira, J.; Nitti, G.; Polo, F.; Lombardero, M.
 Mol. Immunol. 28, 301-309, 1991
 A:Title: Studies on the biochemical structure of the major cat allergen Felis domesti
 A:Reference number: A53283; MUID:91287714
 A:Accession: B53283
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <DUF>

Query Match 6.3%; Score 7; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3,7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 PLYDVF 33
 |||||

DB 8 PLYDVF 14

RESULT 15
 T20911
 hypothetical protein P14F0.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20911
 R:Lloyd, C.
 submitted to the EMBL Data Library, March 1997

A:Reference number: 219346
A:Accession: T20911
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-73 <MIL>
A:Cross-references: EMBL:Z92782; PIDN:CAB07183.1; GSPDB:GNO0023; CESP:F14F8.8
A:Experimental source: clone F14F8
C:Genetics:
A:Gene: CESP:F14F8.8
A:Map position: 5
A:Introns: 16/3; 46/3

Query Match 6.3%; Score 7; DB 2; Length 73;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 LVLALL 13
|||||
DB 5 LVLALL 11

Search completed: May 7, 2002, 12:18:15
Job time: 111 sec

THIS PAGE BLANK (USPTO)
